

MONTHLY NOTICES
OF THE
ROYAL ASTRONOMICAL SOCIETY

Vol. 106

No. 3

MEETING OF 1946 APRIL 12

Professor H. H. Plaskett, President, in the Chair

The election by the Council of the following Fellows was duly confirmed:—

Norman Wilson Allison, B.Sc., 439 Leasons Road, Moreton, Wirral, Cheshire;

Mahmoud Khairy Mohamed Aly, Solar Physics Observatory, Cambridge;

Noel Atherton, 5 Kipling Avenue, Bath, Somerset;

Gerald M. Clemence, U.S. Naval Observatory, Washington 25, D.C., U.S.A.;

George Henry Corbett, 34 Butleigh Avenue, Victoria Park, Cardiff, Glamorganshire;

John Logan Lewis, B.A., The College, Eastbourne, Sussex;

Edward Thomas Pierce, Solar Physics Observatory, Cambridge;

John Robert Pikett, 93 Whitmore Road, Harrow, Middlesex;

William Hugh Cowie Romanis, M.A., M.Ch., F.R.C.S., 149 Harley Street, London, W. 1;

John M. C. Scott, M.A., 37 Church Street, Malvern, Worcestershire;

James Willoughby Somers, B.Sc., 37 Couston Street, Dunfermline, Fife;

Harry Verdon Stopes-Roe, Norbury Park, Dorking, Surrey;

Frank Tyler, B.Sc., Ph.D., 47 Colenso Road, Blackburn, Lancashire; and

Thomas James Willmore, M.Sc., Ph.D., Research and Development Establishment, Cardington, Bedford.

The following candidates were proposed for election as Fellows of the Society, the names of the proposers from personal knowledge being appended:—

Ira S. Bowen, Director, Mount Wilson Observatory, Pasadena 4, Cal., U.S.A.
(proposed by A. H. Joy);

Michael Edwin Chivers, c/o the Authors' Club, 2 Whitehall Court, London, S.W. 1 (proposed by M. Davidson);

Ernest William Mason, Beechwood, 1 Osborne Road, Southampton, Hampshire (proposed by W. H. Steavenson);

Isaac Eyton Roberts, B.Sc., Eyton House, Bagillt, Flintshire (proposed by J. F. Cox);

- Hugh Simmons, Solaris, Edlesborough, near Dunstable, Bedfordshire (proposed by E. L. Hawke);
Horace Geoffrey Spalton, Castle Fields, Willington, Derbyshire (proposed by L. J. Comrie); and
Wladyslaw Jerzy Swiatecki, 170 Victoria Road, Ruislip, Middlesex (proposed by H. Dingle).

Forty-six presents were announced as having been received since the last meeting.

The President called upon Sir Harold Spencer Jones, Astronomer Royal, President of the International Astronomical Union, to give some account of the recent meeting of the Executive Committee of the Union at Copenhagen. The Astronomer Royal spoke of the many problems confronting the Union which could not be deferred until the next general assembly and explained that the Executive Committee had invited representatives of all the member countries to assist it in its deliberations. In particular he referred to the arrangements to be made for the continuance of the work—variable stars, ephemerides, minor planets, bibliographies, etc.—hitherto done by German institutions. He further mentioned the setting-up of a Commission, in co-operation with the United Nations Educational, Scientific and Cultural Organization, to facilitate the interchange of astronomers between the observatories in different countries, and the project of an International Observatory. It was not considered possible to arrange a general assembly of the Union in 1947, but preliminary arrangements were being made to hold such a meeting in 1948.

SOME PHENOMENA CONNECTED WITH INTERSTELLAR MATTER

George Darwin Lecture, delivered by Professor J. H. Oort on 1946 May 10

The first characteristic to strike the investigator of interstellar material is its uneven distribution. It strikes him in the weird shape of dark nebulae, in the chaotic luminous nebulae like that in Orion. It has been found again, by more subtle investigations, in the character of interstellar lines. They indicated that the interstellar gas consists of clouds or currents with considerable relative motions. These motions were first found by Beals* and by Sanford†, who showed that in some stars the interstellar lines were double or even multiple. Quite recently, from his beautiful Coudé spectrograms, Adams‡ has been able to prove that multiplicity is a general feature of interstellar lines.

Previously, O. C. Wilson and Merrill§, stimulated by a paper by Sir Arthur Eddington||, had concluded from a study of the intensity ratio of the interstellar D-lines, combined with a general study of the relation between intensities of interstellar lines with distance and galactic longitude, that the average peculiar velocity of interstellar streams of gas was of the order of 15 to 20 km./sec.

It is tempting to assume that the streams of gas which give rise to the displaced interstellar components are to be identified with the objects observed as dark or luminous nebulae. As a matter of fact, we know from their spectra that several of the diffuse nebulae are gas clouds. The dark clouds must generally owe their darkness to small solid particles, but it is likely that not only solid material but gas as well is concentrated in these clouds. This at least seems to be intimated by the frequent association of luminous gaseous nebulae with dark clouds, and in particular by the phenomenon of the prominent emission edges observed along several dark nebulae. It is also supported by Hubble's¶ classical observation that nebulae excited by stars of type O and B₀ show always gas-spectrum; if concentrations of solid particles and concentrations of gas formed separate objects there is no reason why an O-type star should not sometimes also illuminate a cloud mainly of solid particles. The strongest argument for the identity of the two kinds of clouds is that the gas will largely carry along the solid particles embedded in it, so that, by whatever process the gas has been concentrated into moving clouds, the solid particles will have been similarly concentrated and set into motion. The importance of the interaction between gas and interstellar solid particles was first stressed by Spitzer.**

It has been objected that if gas clouds and clouds of solid particles were identical, stars shining through dense absorbing clouds should show a corresponding strengthening of interstellar absorption lines in their spectra, which has not been observed. If the intensities of the absorption lines due to a cloud were proportional to the number of atoms in the cloud these negative observations would indeed prove that the cloud could not contain a proportionate concentration

* *M.N.*, 96, 661, 1936.

† *P.A.S.P.*, 51, 238, 1939.

‡ *Ap. J.*, 97, 105, 1943; (*Contr. Mt. Wilson Obs.*, No. 673).

§ *Ap. J.*, 86, 44, 1937; (*Contr. Mt. Wilson Obs.*, No. 570).

|| *M.N.*, 95, 2, 1934.

¶ *Ap. J.*, 56, 162, 1922; (*Contr. Mt. Wilson Obs.*, No. 241).

** *Ap. J.*, 93, 369 and 94, 232, 1941.

of gas. But observations of the intensity ratio of the D-lines of sodium show that the intensity is far from proportional to the number of atoms. O. C. Wilson and Merrill have pointed out* that the observed ratio of this doublet is in general about 1.2, instead of 2, as it would have to be in case of simple proportionality. Under these circumstances a sensible strengthening of interstellar lines by a dense interstellar gas cloud is not to be expected, so that there is no longer any observational objection to supposing that the dark clouds contain proportionate concentrations of gas. On the other hand it has been shown that, on the average, the intensities of interstellar lines increase proportionally with the distances of the stars, and that they do so in all galactic longitudes. The only possible explanation of this appears to be that these lines are mainly formed in clouds having considerable "peculiar" velocities, averaging 15 or 20 km./sec. in the radial component.

These considerations also explain why so little general correlation has been observed between interstellar line-intensities and colour-excesses. A *negative* correlation would even be conceivable. For it is quite evident that the distribution of interstellar material is most uneven. Suppose, for a moment, that it would exclusively occur in fairly dense clouds. Among stars of a given apparent as well as absolute magnitude, those most strongly dimmed by absorption, and thus most reddened, should then be the nearest, and would therefore show only weak interstellar lines.

It has been thought, for a time, that solid matter was more strongly concentrated to the galactic plane than the gas, but recent discussions by van Rhijn have tended to show that the differences found in earlier investigations were not real.†

Although there are thus strong reasons for admitting that gas clouds and clouds of solid particles are generally the same objects, it should be pointed out that there appear to exist differences in the *details* of the distribution of gas and particles, such as have been indicated by Struve in the diffuse luminous nebula around σ Scorpii.‡

An important question that may be asked is whether there is a connection between the distribution of early-type stars and that of interstellar material. I have not studied this problem in detail, but I do not think the available evidence points to any general connection. The luminous nebulae seem to be nothing but parts of the general system of interstellar clouds that happen to lie close to a star of high luminosity. There may, however, be a relation in a few special cases, like, for instance, the Orion group, where a clustering of early-type stars coincides with a pronounced concentration of interstellar clouds.

Because of their very uneven character, it is difficult to give more than an extremely schematic description of structure, dimensions and numbers of interstellar clouds.

In our neighbourhood are found two extended systems of dark clouds, one in Taurus-Auriga-Perseus, the other in Ophiuchus and Scorpius. According to Pannekoek's counts of the Bonner Durchmusterung§ both systems cover regions of about 500 square degrees and show average absorptions of $0^m.8$. Similar

* *Ap. J.*, 86, 44, 1937; (*Contr. Mt. Wilson. Obs.*, No. 570).

† *Publ. astr. Lab. Groningen*, No. 50, 1946.

‡ *Ap. J.*, 86, 94, 1937; (*Contr. McDonald Obs.*, No. 5).

§ *Publ. astr. Inst. Univ. Amst.*, 1, 1924.

results have been derived more recently from the star-counts carried out in the United States.* As they lie at a distance of between 100 and 150 parsecs their outer diameters are of the order of 50 parsecs, while the denser nuclei may measure 30 parsecs. As is well known from the beautiful photographs by Ross and others, from the star-counts by Dyson and Melotte†, as well as from numerous later investigations, the distribution of the interstellar material within these large cloud complexes is again most uneven. Practically clear spaces are interspersed with regions of strong absorption; in some very dark clouds, with diameters of a few parsecs, the density appears to be as much as 1000 times the average interstellar density.

A still finer structure is revealed where the clouds are illuminated by stars, as in the region of the Pleiades, where large-scale photographs of the region around Merope show a set of regularly spaced streaks, perhaps half a parsec long, and no more than about 1000 astronomical units wide.

Very interesting information on details of interstellar distribution was obtained when the light-signal from the outburst of Nova Persei 1901 slid along the wisps of interstellar matter in front of this star, and gave rise to the rapidly expanding luminous ring sectors that were observed during the first two years after the outburst. The points that are simultaneously seen illuminated by the nova signal are situated on the surface of an ellipsoid of revolution with focal points in the nova and the Sun. In 1934 Humason's measures of the expansion of the small gaseous envelope which became visible around Nova Persei in 1917, showed that the distance of the nova was about 500 parsecs. At this distance the observed expansions in the outer rings observed in 1901 and 1902 correspond to velocities many times larger than that of light. This, at first sight, somewhat surprising result is easily explained when we realize that the luminous rings are situated at considerable distances, up to 10 parsecs, in front of the nova, where the velocity of expansion of the locus of points seen in the light of the nova at a given instant exceeds considerably the velocity of light. This nova outburst was the only occasion which permitted us to obtain a three-dimensional picture of the distribution of interstellar particles. The material in this region appears to be mainly distributed in two or three thin layers from $1/20$ to possibly $1/2$ parsec thickness, and extending from 5 to 10 parsecs in the other dimensions. Very remarkable is the distribution of interstellar matter immediately behind and beside the nova. This is highly unsymmetrical and has a queer fan shape, the top of the fan pointing to the nova.‡

We cannot make a reliable estimate of the extent to which interstellar material is concentrated in clouds, but it is useful to try a rough guess.

Of the 160 stars with spectral types between O and B2 which are brighter than $7^m.5$ and north of -15° , 23 occur in Hubble's list of stars connected with diffuse nebulosity. If we assume that, in general, there is no relation between interstellar clouds and early-type stars we may conclude that the cloud regions occupy 15 per cent. of the space near the galactic plane. Now if we conclude further, from the observations on the multiplicity and intensity of interstellar lines, that stars at a

* McCuskey, *Ap. J.*, **88**, 209, 1938; **89**, 568, 1939, and **94**, 468, 1941.

† *M.N.*, **80**, 3, 1919.

‡ Cf. Ritchey, *Ap. J.*, **15**, 129, 1902. The distribution of interstellar material around Nova Persei 1901 and the phenomena connected with the expanding light-signal were independently studied in 1939 by the author and Mr de Kort at Leyden, and by P. Couderec. The latter published his investigations in *Ann. Astrophys.*, **2**, No. 2, 1939.

distance of 1000 parsecs shine on the average through five different cloud regions (which seems conservative) we arrive at an average radius of 20 parsecs for each cloud region, and an average distance apart of 65 parsecs.

It is probable from the observations of interstellar lines, as well as from studies of absorption, that a considerable fraction of the interstellar material is concentrated in these clouds. But it is still impossible to make a significant statement on the density in the space between the cloud systems: I do not think that present-day observations would permit us to decide between whether it is, say, one-third, or one-hundredth of the general average interstellar density.

We have seen that the cloud regions move with peculiar velocities of the order of 15 km./sec. in one co-ordinate. As there appears to be no system in these motions they would seem inevitably to lead to encounters between clouds. With the dimensions and distances just estimated a given cloud system would once in 10^7 years meet with another system in such a way that the relative velocity of the two systems would be largely eliminated. By such encounters the individual cloud systems and their relative motions would disappear in a time short compared to the probable age of the galactic system. As they have not disappeared a mechanism must be at work by which they are being reformed and set into motion.

An entirely different group of data seems also to bear witness to the continual dissolution and reformation of interstellar clouds. Observations of the last fifteen years, and particularly the six-colour photometry recently carried out by Stebbins and Whitford*, have shown that the dependence of interstellar scattering upon wave-length is accurately the same for all stars†; the same therefore, whether we observe through a dense cloud, like that in Taurus, or through a thin medium without any evident concentrations. This means that particles within dense clouds have the same radii as outside particles. If we accept as working hypothesis that these particles grow by condensation from the interstellar gas, this curious result can only be comprehended if the clouds are being continually reshuffled.

The mechanism by which the formation of the clouds is brought about is still obscure. Possibly, the ultimate cause of the erratic motions and forms of interstellar clouds might be found in the differential galactic rotation, by which the layers of interstellar gas at different distances from the centre are forced to move with different angular velocities. As Reynolds's number is of the order of 10^9 for our region of the galactic system, turbulent motion might be expected to result.‡ Professor J. M. Burgers of Delft, to whom I am much indebted for various discussions on this aspect of my subject, has indicated that it is not unlikely that under these circumstances the units of the turbulent motions might be of sizes comparable with the separations between interstellar clouds. On the other hand, one must keep in mind that in several essential points the arrangement and motion of interstellar material differs radically from that of ordinary gas in turbulent motion.

We have considered the smallest observable details of the inner structure of interstellar clouds. We have also considered the arrangement of interstellar material in clouds and systems of clouds. We may finally ask whether there is a

* *Ap. J.*, **98**, 20, 1943; (*Contr. Mt. Wilson Obs.*, No. 680).

† An exception is formed by the stars in the Orion nebula.

‡ Zwicky has been the first to draw attention to the fact that Reynolds's number is generally large for interstellar media. Cf. *Ap. J.*, **93**, 411, 1941.

still larger-scale arrangement of the dark material in the galactic system. Evidence of such an arrangement may perhaps be found in the well-known dark rift which apparently divides the Milky Way in Aquila into two Ways. More remarkable evidence comes from the Cepheids in the Cygnus cloud. In the brightest part of this star cloud there are some faint Cepheids discovered by Baade. Colours measured by Oosterhoff show them to have slight or no colour-excesses. The absorption in this direction, as inferred from these colour-excesses, appears to be exceptionally small, and the corresponding distances of the stars surpass greatly any other distances observed for objects near the galactic plane; they run up to about 20,000 parsecs, the farthest being twice as far from the centre of the Galactic System as the Sun. As shown in Table I the radial velocities of three of

TABLE I
Distances and velocities of Cepheids in the Cygnus cloud

	r	ϖ	Θ_c	v_c	v_0
Sun	0.0	10.0	260	...	...
GL Cygni	7.9	11.4	246	- 43	- 41
V 343 Cygni	13.2	14.8	212	- 114	- 97
QY Cygni	12.0	13.7	223	- 94	- 112
V 336 Cygni	24.0	23.8	122	- 204	- 107

(r is the distance from the Sun, ϖ that from the centre of the Galactic System, both in kps. The column marked Θ_c indicates the estimated rotational velocity of the Galactic System, while v_c and v_0 give, respectively, the computed and observed radial velocities in km./sec., corrected for the standard solar motion. The average galactic co-ordinates of the Cepheids are $l = 42^\circ.3$, $b = +3^\circ.5$)

these objects determined by Joy* agree quite satisfactorily with the motions to be expected from the galactic rotation at these distances†, so that it is very probable that the distances are of the right order. The situation of the Cepheids relative to

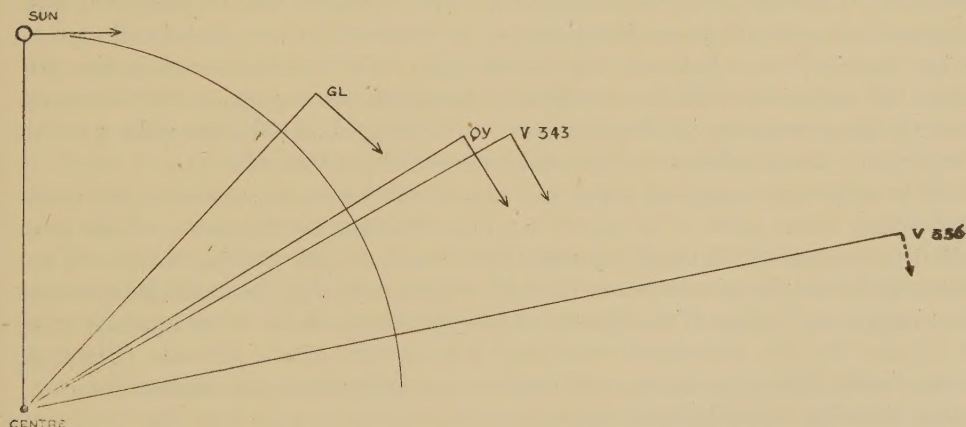


FIG. 1.—Cepheids in Cygnus cloud.
The arrows indicate the rotational velocity of the Galactic System.

the Sun and the galactic centre is indicated in Fig. 1. We find thus that a stretch of at least 10,000 parsecs close to the galactic plane contains at most as much interstellar material as is normally contained in a column of 1000 parsecs. This can

* *Ap. J.*, 86, 363, 1937; (*Contr. Mt. Wilson Obs.*, No. 578), and private communications.

† *B.A.N.*, 9, 325, 1942.

hardly be due to chance, but points to a large-scale arrangement of interstellar clouds, by which a long lane is kept practically empty.

Very unexpected results are shown by counts of faint extra-galactic nebulae in this region, which seem to stand in sharp contrast with the conclusions just stated. Baade writes that careful counts to a very faint limit show only a few nebulae in the region of the Cepheids, indicating that out beyond this large distance from the galactic centre, and also beyond 1400 parsecs from the galactic plane, there must still be surprisingly large amounts of absorbing material.

We now return to the random motions of the interstellar clouds, and to the resulting encounters. An encounter between two clouds will be accompanied by two effects which I wish to consider this afternoon.

At a head-on encounter between two clouds a compressed intermediate region will be formed. If the densities and temperatures of the two clouds are supposed to be equal, and their relative velocity 40 km./sec., both the density and the temperature in the intermediate layer will become three times that in the original clouds. The compressed region will be at rest relative to the average motion of the two clouds. The transition between the clouds and the compressed layer will be quite abrupt. The numbers quoted should be representative for encounters between clouds in general. It may be noted that collisions between interstellar clouds will, as a rule, occur with velocities exceeding that of sound, so that we are dealing with shock-waves. I am again indebted to Professor Burgers for information on these points.*

The rise in temperature in the intermediate layer will easily be sufficient to cause hydrogen to be excited and ionized by collisions, and the gas will therefore become luminous. In case the two clouds have different densities, the densities in the compressed layers formed in the two clouds will differ in the same ratio as the original densities; the temperatures will be inversely proportional to these densities. Consequently, the compressed layer formed in the thinner of the two clouds will in general become luminous, while that in the denser cloud will remain at low temperature. It is not improbable that at the interface between the two layers the luminosity will be considerably enhanced on account of the increased density, the accompanying drop in temperature being of no influence as long as the temperature remains above a certain minimum (of about 15,000°).

It is tempting to suppose that it is this strongly luminous interface, and more particularly those parts of it which are perpendicular to the plane of the sky, which causes the bright rims frequently observed near dark nebulae when these are bordered by luminous nebulae, and that the adjacent thin luminous fields represent the compressed regions of the thinner of two colliding clouds. The luminescence of nebulae like the well-known extensive nebulosities near ζ Orionis (Plate 5), or the North America nebula, may well be caused by this collisional mechanism, rather than by radiation from stars.

An elementary calculation shows that the amount of energy set free by an encounter as considered could easily be of the right order to make up for the total energy radiated by such a nebula. Other diffuse nebulae which are not evidently excited by a star may shine by the same agency.

A second effect to which I want to draw attention is the evaporation of the solid particles of the two colliding clouds when the particles of one of the clouds, upon

* *Proc. Akad. Sci. Amst.*, 49, 589, 1946.

penetrating into the intermediate layer, hit other solid particles of comparable sizes. It can be shown that the probability of such an evaporation is by no means negligible, and that it may play an important part in the household of interstellar material.

I cannot enter here upon a third phenomenon which apparently may accompany an encounter between two clouds, viz. the wavy structure that is so beautifully displayed in the nebulosity surrounding Merope.

The effect of evaporation of particles brings us to the problem of the origin and growth of the solid interstellar particles.

It is now generally agreed that the absorption or scattering of light in the Galactic System must mainly be due to particles with radii of the order of 1000 angstrom units. There is difference of opinion regarding the composition of these particles, whether they are absorbing metallic, or non-absorbing non-metallic. There is also difference of opinion concerning their origin. Some look upon them as meteoric "dust", others as condensations in the interstellar gas. Though certain observational data seem unfavourable for the conception of metallic particles, they cannot be definitely ruled out on these grounds. Nor are there any observations which could distinguish between meteoric particles and particles condensed from the interstellar gas. However, the considerations given below make a meteoric origin of the scattering particles improbable, inasmuch as it is difficult to see how even the maximum number of meteors that might be admitted could yield dust at a rate sufficient to explain the interstellar scattering.

In any case it seemed of interest to investigate more closely the possibility of condensation in the interstellar gas. The attack on this problem was opened in 1935 by Lindblad*, who showed that, once nuclei had been formed, particles of appropriate size might grow by sublimation of interstellar gas around these nuclei. In 1941 the University of Leyden, though at that time already almost entirely closed by the Germans, offered a prize for a discussion on the formation of nuclei, and on the distribution of diameters of particles originated in the interstellar gas. The essays sent into the University formed the starting-point for some more extensive discussions, among others those of Kramers and ter Haar on the formation of molecules, and of nuclei for condensation†, and a study by Mr van de Hulst and myself on the process of evaporation, the distribution of radii and the resulting law of scattering.‡ In the following section of my lecture I shall report on this latter investigation.

Though it appears still impossible to compute the actual number of nuclei which will be formed, Kramers and ter Haar have shown that it is quite plausible that this number would be of the right order to explain the observed numbers of particles.

I shall assume as a working hypothesis that the interstellar particles were formed that way, and that they grow by sublimation of gas atoms on to the particles, practically every atom or ion hitting the cold particles sticking to it. The process being analogous to that by which smoke is formed, we propose to designate the solid interstellar particles as smoke.

* *Nature*, **135**, 133, 1935.

† Ter Haar, *B.A.N.*, **10**, 1, 1943 (also *Ap. J.*, **100**, 288, 1944); Kramers and ter Haar, *B.A.N.*, **10**, 137, 1946.

‡ *B.A.N.*, **10**, 187, 1946. Cf. also *Ned. Tijdschr. Natuurk.*, **10**, 238, 1943, and *Versl. Akad. Sci. Amst.*, **55**, 18, 1946.

Upon closer inspection it becomes apparent that this picture of the one-sided process of condensation of the gas onto the solid particles, without occurrence of an inverse process, cannot be complete. The nature of the difficulty is this. If the smoke has originated by condensation in the interstellar gas the speed of its formation must increase with at least the third, and probably with a rather higher power of the density. The number of solid particles present after a certain time will therefore be strongly dependent upon the original gas density ρ_0 . If for a certain value of ρ_0 the present densities of smoke and condensable gas (that is, all gases other than hydrogen and helium; van de Hulst having shown that under interstellar conditions particles made up of hydrogen and helium would evaporate) would be comparable, a relatively slight increase in ρ_0 would have caused a practically complete disappearance of all condensable gases, while for smaller ρ_0 we would soon come to a point where no appreciable amount of solid particles could have been formed. In most places we should thus expect either no solid particles to have formed at all, or else the condensation to have practically exhausted the condensable gases. The intermediate phase, where these gases and the solid particles would exist side by side, should be of rare occurrence. This is contrary to observations. For not only do we find this very condition in our surroundings, and in all observable parts of the galactic system, but it is similarly found in all extra-galactic systems in which absorption is found, and at the same time conditions are such that gas atoms can be excited.

We are thus forced to conclude that one-sided condensation cannot account for the observed data, and that a process must be at work by which gas is being released from the solid particles.

I should mention that I was made to think about this problem by reading the prize essay sent in to Leyden University by Mr van de Hulst.

It has been shown that at the low temperature of the smoke particles ordinary evaporation cannot be of any importance except for hydrogen and helium. Photo-dissociation of atoms from the surface might be considered. From theoretical discussions it does not seem likely that this would play an important part. In the following discussions I shall tentatively discard it. The only other process by which gas can be set free is by total evaporation of smoke particles as a consequence of a collision between two such particles. If we compute the frequency of collisions from the data on the frequency, sizes and motions of interstellar clouds as estimated above, we find that the evaporations through collisions must be quite important, and that they may well account entirely for the amount of gas present in interstellar space.

Under these circumstances it seemed of interest to investigate the state which will result after conditions in interstellar space have stabilized; that is, when smoke particles have grown in such numbers and to such sizes that further birth and growth are just balanced by evaporation through collision. This state will then be confronted with observation, namely with the observed average coefficient of absorption, and with the way in which the absorption varies with the wave-length.

If $n(r)dr$ represents the number of particles with radii between r and $r + dr$ the condition for a stationary state is

$$\dot{r}n(r) = \dot{r}n(r + dr) + P(r)n(r)dr. \quad (\text{I})$$

Here $\dot{r}$ is the growth of the radius per unit time, $P(r)$ is the probability of evaporation in a unit of time. The left-hand member gives the number of particles

growing into the interval from r to $r + dr$, the first term in the right-hand member shows the number leaving it through growth, the second term shows the number of evaporations. For simplification the melting together of the two particles, which will occur when they hit each other with relatively low velocities, has been left out; it can be shown that its inclusion would not have made any important difference in the result. The quantity $\dot{r}$ may be considered as a parameter which is independent of r .

The solution of the equation is complicated by the fact that $P(r)$ depends implicitly upon $n(r)$. When a cloud sweeps up detached interstellar particles these will penetrate into the cloud. The gas in the cloud will gradually diminish their velocities, the distance to which they penetrate being proportional to the mass of the particles. For the particles concerned this distance will in general be small compared with the dimensions of the cloud. The chance of evaporation through collision with another particle will evidently be larger for larger particles. If $n(r)$ were constant it would be proportional to the fourth power of r . It is evident therefore that this process will tend to eliminate the larger particles, and that above a certain r the number of particles must show a very rapid decrease.

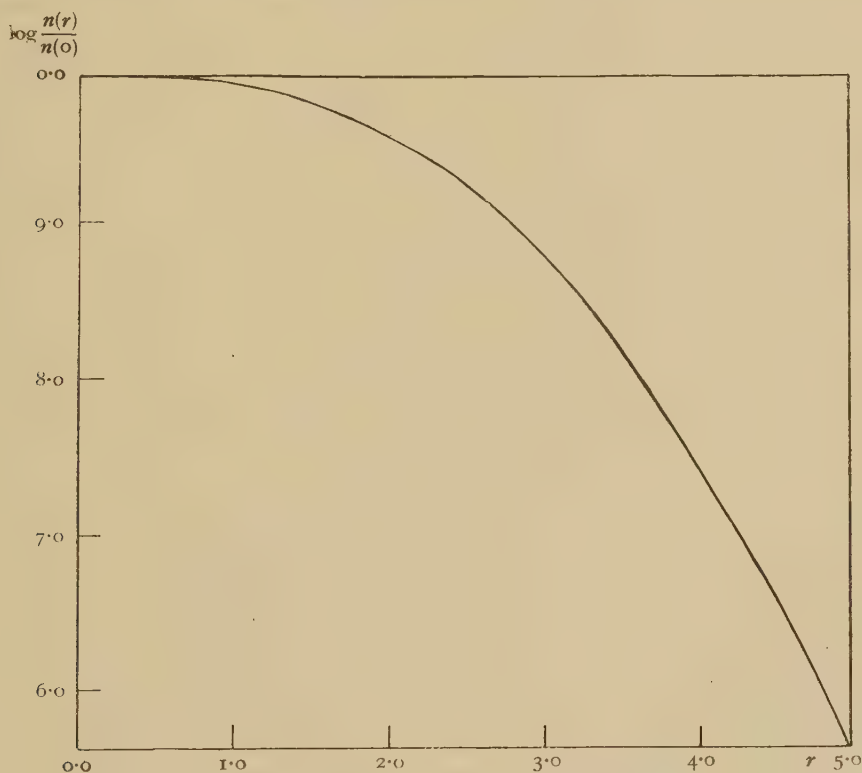


FIG. 2.—Distribution function of radii of interstellar particles.
The unit in which the abscissæ are measured is about 0.134 microns.

The curve in Fig. 2, showing $\log n(r)/n(0)$ as a function of r , was obtained by numerical integration [$n(0)$ is the number of particles per unit interval for $r=0$]. The form of the curve is entirely determined by the differential equation just given; it does not depend on the shapes or distribution of interstellar clouds and other

uncertain data. It is only the *scale* of the abscissæ which depends on these. This unit of length in the figure is given by

$$l = \left\{ \frac{\dot{r}}{kn(0)} \right\}^{1/5}, \quad (2)$$

where k is a quantity measuring the probability that a given cloud encounters another in a unit of time. With the dimensions and motions of clouds as estimated k would be 1.3×10^{25} per million years. It may be noted that the only datum needed for computing k is the average number of clouds through which the light of a star at a given distance passes before it reaches us. No data regarding size or space frequency of clouds are required. The velocity of growth, $\dot{r}$, is determined by the density of the condensable gases, oxygen, carbon, nitrogen, etc. The interstellar density of hydrogen has been estimated as of the order of 2 atoms per cm^3 . If the relative abundances were the same as in the Sun's atmosphere there would be about one heavier atom to a thousand atoms of hydrogen and helium. Relative amounts of the same order seem to be indicated in the Orion nebula.* On the assumption that every "heavier" atom hitting the smoke particle sticks to it, and that the density of the solid particles is unity, $\dot{r}$ would be equal to 4.5×10^{-7} cm./million years . The third parameter, $n(0)$, is determined by the number of new particles that are formed per cm^3 and per unit of time. If this is called N , we have $n(0) = N/\dot{r}$. From the observed density of interstellar molecules it can be inferred that $n(0)$ is smaller than 10^{-3} ; the theory of the formation of larger molecules has not yet been developed sufficiently to tell how much smaller, but it is likely to be several factors of 10.

If we want to compare the present theory with observations, i. e. with data on the average absorption of light and on the law of interstellar scattering, we shall have to make our computations for various values of $n(0)$. We have carried out calculations for $n(0) = 10^{-9}$, 10^{-7} , 10^{-5} and 10^{-3} , respectively. We have also varied the other parameters, k and $\dot{r}$, in the way shown in Table II. These latter variations must be considered as extreme possible variations, the actual values may be expected to be near the average values of 1.3×10^{25} and 4×10^{-7} mentioned above.

The columns $r_{1/2}$ show the radii for which the frequency has dropped to one half of the frequency for the smallest radii. This may be taken to represent roughly the point where the frequency curve begins its steep drop. The values of $r_{1/2}$ are seen to vary from a few tenths to a few hundredths of a micron. The variation is relatively small compared to that of the parameters, as a consequence of the fifth root in (2).

Assuming that the particles are non-conductive (as they would be on the present theory of their formation) and have a refractive index $4/3$, like ice, we can now calculate how much the particles of various sizes will contribute to the scattering. The scattering coefficient being smaller for the smaller particles the total scattering will first increase with r ; at a certain value of r it will begin to decrease again, because of the drop in the number of particles. The radius for which the scattering is a maximum may be called the effective radius; it is shown in the last column of the table. It is of the same order as $r_{1/2}$. The scattering coefficients

* Wyse, *Ap. J.*, **95**, 356, 1942; (*Lick Obs. Contr.*, Ser. II, No. 4); Aller and Menzel, *Ap. J.*, **102**, 239, 1945.

TABLE II

Constants of the frequency function of the radii of intersellar particles, photographic absorption coefficients, and effective mean radii for various combinations of the parameters $n(o)$, k and $\dot{r}$

$n(o)$	$\dot{r} \backslash k$	$4 \cdot 10^{-8}$	$4 \cdot 10^{-7}$	$4 \cdot 10^{-6}$	$4 \cdot 10^{-8}$	$4 \cdot 10^{-7}$	$4 \cdot 10^{-6}$	$4 \cdot 10^{-8}$	$4 \cdot 10^{-7}$	$4 \cdot 10^{-6}$
		$r_{1/2}$ (in microns)			$\log a_{pg}$			r_{eff} (in microns)		
10^{-9}	$1 \cdot 3 \cdot 10^{24}$	.59	.94	1.49	0.56	1.10	1.65	.63	.93	1.38
	$1 \cdot 3 \cdot 10^{25}$	.37	.59	.94	9.92-10	0.56	1.10	.42	.63	.93
	$1 \cdot 3 \cdot 10^{26}$	.24	.37	.59	9.05-10	9.92-10	0.56	.29	.42	.63
10^{-7}	$1 \cdot 3 \cdot 10^{24}$	.24	.37	.59	1.05	1.92	2.57	.29	.42	.63
	$1 \cdot 3 \cdot 10^{25}$	.15	.24	.37	9.99-10	1.05	1.92	.19	.29	.42
	$1 \cdot 3 \cdot 10^{26}$	.09	.15	.24	8.83-10	9.99-10	1.05	.13	.19	.29
10^{-5}	$1 \cdot 3 \cdot 10^{24}$	.09	.15	.24	0.83	1.99	3.05	.13	.19	.29
	$1 \cdot 3 \cdot 10^{25}$	.06	.09	.15	9.64-10	0.83	1.99	.08	.13	.19
	$1 \cdot 3 \cdot 10^{26}$	.04	.06	.09	8.41-10	9.64-10	0.83	.05	.08	.13
10^{-3}	$1 \cdot 3 \cdot 10^{24}$	.04	.06	.09	0.41	1.64	2.83	.05	.08	.13
	$1 \cdot 3 \cdot 10^{25}$	.025	.04	.06	9.13-10	0.41	1.64	.035	.05	.08
	$1 \cdot 3 \cdot 10^{26}$	.015	.025	.04	7.78-10	9.13-10	0.41	.020	.035	.05

used were taken from an investigation by Greenstein.* More extensive calculations have now been carried out by van de Hulst.† It may be noted at once, that the effective radii found are at least of the right order to explain the law of scattering observed in interstellar space. This must, I think, be considered a remarkable result in itself, because the law of scattering has in no way entered into the theory or into the constants considered so far.

We now proceed to consider the average photographic absorption coefficients, which may at once be calculated from the law of scattering and the distribution function of the radii as determined by each combination of the parameters. The results are shown under $\log a_{pg}$, a_{pg} being the absorption in magnitudes per kiloparsec. These should be confronted with the observed average absorption coefficient, which may be taken as $1^{m.5}$ per kps., or $\log a_{pg}=0.2$. We observe that there is general agreement in the order of the two absorptions, especially if we limit ourselves to values of $\dot{r}$ and k close to the most probable values (the quantities corresponding to these most probable values have been printed in bold face type). This agreement is again quite remarkable, the theoretical absorption coefficients might just as well have come out a million times larger or smaller. To some extent the result might thus be considered to speak for the present theory. We may note, in passing, that the absorptions are in no way proportional to the numbers of particles of small radii: between $n(o)=10^{-3}$ and $n(o)=10^{-7}$, a decrease in $n(o)$ appears even to entail an increase of the absorption, the ten thousand-fold decrease in $n(o)$ corresponding, in the domain considered, with about a five-fold increase in a_{pg} .

It is now of interest to reverse the problem, and to inquire for which values of

* *Harvard Circ.*, No. 422, 1937.
† Thesis, Utrecht, 1946.

the three parameters the observed absorption and its dependence upon the wavelength can be reproduced. We can determine only two unknowns from these conditions, and will therefore assume that $\dot{r} = 4 \times 10^{-7}$. The observations may then roughly be satisfied by taking $k = 6 \times 10^{25}$, $n(0) = 10^{-7}$.^{*} The number of solid particles per cm.³ would then be of the order of 10^{-12} .

The smallness of $n(0)$ would indicate that there are a considerable number of "difficult" steps in the building up of the nuclei of the smoke particles. It shows that the formation of smoke particles would be proportional to a high power of the gas density, of the order of ρ^8 .

This result leads us again back to a remark from which the present discussion started: If the formation of smoke particles depends upon so high a power of the original gas density, how can it then be that different galaxies, in which the gas density will certainly not be exactly equal, exhibit absorption effects of more or less the same order, and show also similar present densities of interstellar gas? The answer is that indeed in a system in which the gas density was 10 times that in the Galactic System 10^8 as many particles will be formed. But they do not get a chance of growing, as the probability for a particle to vaporize through collision grows proportionally. The following table shows the absorption coefficient which would correspond to different values of the gas density, ρ . The value $\rho_g = 3 \times 10^{-24}$ represents the density in our neighbourhood. The computations were made on the supposition that k is the same in all cases.

TABLE III

Estimated absorption coefficients and effective radii of interstellar particles for different values of the original gas density

ρ	ρ/ρ_g	a_{pg} (mag./kps.)	r_{eff} (microns)
0.3×10^{-24}	0.1	0.0001	0.5
0.9 "	0.3	0.05	1.1
3 "	1	2.0	0.3
30 "	10	1.0	0.03
300 "	100	0.02	0.01

The results in the table show that only for a certain range of density, roughly between 2 and 50 times 10^{-24} gm./cm.³, interstellar particles will be at the same time sufficiently numerous and sufficiently large to give observable absorption effects. If the density falls below 1/3 that in our surroundings the absorption will already vanish. For densities ten times that in our surroundings the absorption will be of the same order as in the galactic system, but the effective radii of the particles will be ten times smaller, and the scattering will be according to Rayleigh's law. For still larger densities the absorption becomes again negligible.

Keeping well in mind that conditions in other stellar systems may not be sufficiently similar to those on which the calculations for this table were based, it is interesting to compare the results found with observations of absorption in other galaxies.

In the first place, the absence of conspicuous absorption in practically all elliptical nebulae may easily be understood. We have only to assume that the gas density in these nebulae is 1/3 of that near the Sun, or less. The alternative of a much greater density, though attractive because it might explain why the mass

^{*} More accurate values have been derived in *B.A.N.*, 10, 203, 1946. These correspond to a unit of length in Fig. 2 of 1.34×10^{-15} cm.

density in a system like NGC 3115 was found so much higher than was expected from the light density, should, I think, be excluded because absorbing rings should then be observed near the rims of these nebulae, where the gas density drops to lower values. This consideration leads to the interesting conclusion that the high mass density found in NGC 3115* is made up of extremely faint dwarfs, which compared to the solar-type stars should be relatively at least 10,000 times as numerous as in the galactic system. The interstellar gas in NGC 3115 would appear to be practically completely exhausted in the formation of the stars. One might venture to suggest that the absence of highly luminous stars in elliptical nebulae is connected with the absence of interstellar gas and smoke.

In the spiral nebulae of the so-called Sa-type, which have been considered as more or less adjacent to the more flattened elliptical nebulae like NGC 3115, absorption phenomena have developed in conspicuous strength. In several, perhaps in most, of these nebulae, the absorbing clouds are confined to a ring-shaped region. The finest example is the well-known nebula NGC 4594 (see Plate 6A). The distribution of light in this nebula has been measured at Leyden from photographs taken with the 60-inch reflector of the Mt. Wilson Observatory.

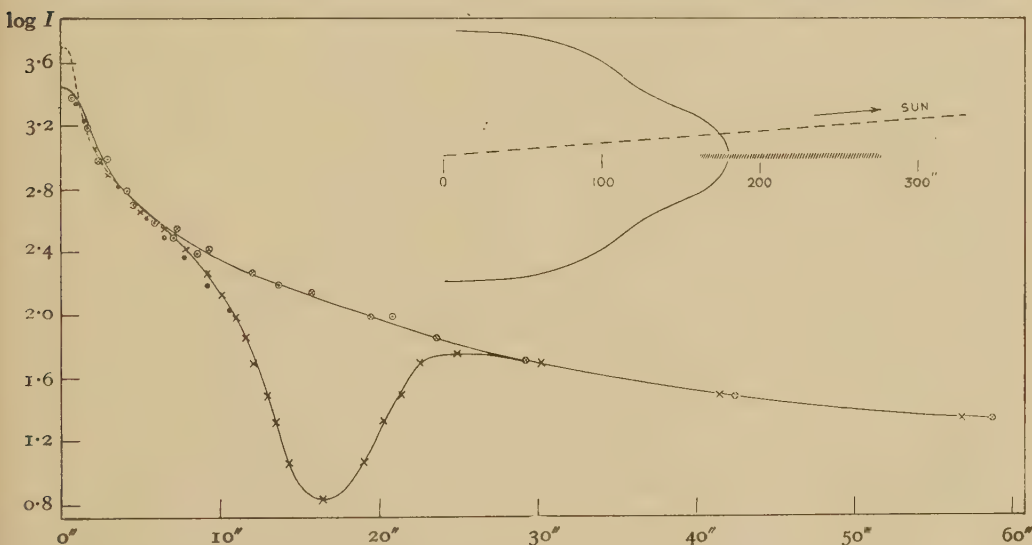


FIG. 3.—Distribution of surface luminosity along minor axis of NGC 4594.

Ordinates are logarithms of surface brightness expressed in stars of $25^{m.0}$ per square second of arc; abscissae are distances from the nucleus in seconds of arc. Measures south of the nucleus, through the shadow band, are represented by simple dots and crosses, the signs referring to measures north of the nucleus have been surrounded by circles. The dashed curve shows the distribution after correction for the light-dispersion on the plates. In the upper right-hand corner a section is shown through the absorbing ring and an outer contour of the luminous system.

Fig. 3 shows the distribution of the light-intensity along the minor axis. The signs surrounded by circles refer to the northern half of the minor axis, the simple dots and crosses to the southern half. It will be seen that there is exact symmetry up to $6''.5$ from the nucleus, indicating that the inner part of the nebula is sensibly free from absorption. The same equality of the northern and southern halves is found beyond $28''$. The absorbing matter is apparently confined to a ring zone sharply bounded at both the inside and the outside. The distance of the nebula

* *Ap. J.*, 91, 273, 1940.

may be estimated at two million parsecs. As the line of sight makes an angle of about 5° with the equatorial plane of the system the ring is found to extend in this plane from 1200 to 2800 ps. from the centre. The ring is a very flat structure, being at most 50 ps. thick. Its inner edge lies only a little inside the outer limit of the luminous nebula as traced on the photographs. The configuration, schematically indicated in Fig. 3, bears some resemblance to Saturn and its rings.

It is tempting to make a comparison between the phenomena in NGC 4594 and the data in Table III, and to put forward the admittedly speculative suggestion that the absence of dark matter in the main central body of the nebula is not due to lack of gas in the inner regions but arises because, increasing inwards, the gas density has in this region grown to values larger than about 100 times that in our surroundings, or larger than the limit of about 3×10^{-22} beyond which no appreciable absorption is to be expected according to Table III.*

Such a large gas density is not incompatible with the spectrographic measures of rotation, which yield an average density of about 1.0×10^{-21} gm./cm.³. The actual value may deviate from this by a factor of at most 3, depending upon the axial ratio of the attracting mass. If a large fraction of this density is made up of gas this would be sufficiently dense to explain the absence of absorbing matter. However, we do not know how large a fraction of the density is due to gas. It is true that if the luminosity law were similar to that found in our surroundings, the density due to stars (calculated from the light of the nebula) would be small compared with the total mass required by the measures of rotation, so that most of this would then have to consist of gas, but the experience gained by the study of the elliptical nebula NGC 3115 warns us not to put much trust in calculations of the stellar density from the light.

But, in my opinion, the suggestion that a large part of the mass would be gaseous receives considerable support from the observed occurrence of spiral arms in the outer parts of the system. For it is hard to conceive how such a structure could have developed if the distribution of mass resembled that of the system of luminous stars, with its strong concentration to the centre and very moderate flattening. In order to have the possibility of an unstable configuration there would have to be a non- or little-luminous substratum of strong flattening, and with a mass at least comparable to that of the stars. Now, exactly such a substratum, composed of gas, and comparable in mass to the entire *stellar* system, is needed for explaining the dark ring on the above hypothesis. The calculations of both Spitzer† and Hoyle‡ have shown that gas, if present in a stellar system, will rapidly cool, and thus concentrate near the equatorial plane.

On the other hand, the absence of dark matter in the elliptical system NGC 3115 has led us to conclude that here the mass of eventual interstellar gas must be

* It may be noted in this connection that, quite apart from the effect of evaporation through collision, there exists a critical gas density such that for larger densities no particles can be formed. This has been first remarked by ter Haar, who showed that at this critical density small particles would be heated up so much by the colliding atoms that they would evaporate. The critical density is estimated to be between 10^{-22} and 10^{-23} (ter Haar, *B.A.N.*, 10, 1, 1943). In an actual case matters are complicated by the fact that particles can be formed at some distance from the equatorial plane, where the gas density is smaller; but it may be shown that ter Haar's mechanism will still prohibit the formation of solid particles in the inner parts of a nebula if the gas density is sufficiently high.

† *Ap. J.*, 95, 329, 1942.

‡ *M.N.*, 105, 287, 1945.

negligible compared with that of the stars. No instability is accordingly to be expected in this case; and indeed, there is no trace of the formation of spiral arms, or of any star clouds or stars of high luminosity which usually accompany spiral structure.

It is noteworthy, in this connection, that in the sequence of nebulae the first development of spiral structure is correlated with the occurrence of a peripheral ring of smoke, and *not at all* with the flattening of the visible systems, no signs of spiral-like structure being visible in any So nebula, though of these many show a most extreme flatness.

I should like to stress that the above suggestions are still quite speculative, and that there are certainly other ways in which the absence of smoke in the inner parts of NGC 4594 might be explained. If I have perhaps dwelt somewhat too long on this speculation it is because this is a phase of the subject in which I have been particularly interested. In fact, one of the motives why the study of the formation of solid particles from interstellar gas was instigated in Holland was that the easily observable smoke in extra-galactic nebulae might teach us something about the *unobservable* gas in these systems. We were interested in the *gas* because this might play a dominant part in determining the field of force.

I should now like to turn to another type of phenomena connected with the interstellar gas. In an earlier part of my lecture I have mentioned that at the encounter of two clouds an intermediate layer is formed in which the temperature may be raised sufficiently to excite and ionize the hydrogen atoms. We might ask if this process could not also be responsible for the luminescence of such nebulae like the Veil or Network nebulae in Cygnus, for which no proper exciting star has been found. Observations show that the Veil nebulae are parts of an expanding shell, possibly that of a supernova. The expansion makes them run up against the interstellar gas, which will raise their temperature.

Before entering upon this special case let us first consider in a more general way what will be the action of the interstellar medium on an expanding shell of gas, like that thrown off by a nova or supernova. I feel considerable hesitation in discussing this subject, inasmuch as my discussion does not reach out above a reconnaissance of the field, and only a first stumbling beginning has been made with the theory of the phenomena concerned. If I have nevertheless risked making public the few results to which these first attempts have led, it is mainly to draw attention to these problems, and to indicate that a more detailed theory might be of value to astronomers. I should also like to say that I have profited very much by discussions with, and help from my friends Professor Kramers of Leyden, and Professor J. M. Burgers of Delft.

If a shell of gas is thrown out from a star which originally lay embedded in the interstellar gas, the latter will be compressed, and to some extent swept up by it. As a consequence, the shell's outward motion will be retarded, its temperature will be raised, and the density distribution in the shell will be changed.

The following formula, first given by Kramers, shows what roughly may become the relation between the number of atoms per cm.³ in the shell, N , that in the interstellar medium, n , the average random velocity of the atoms in the shell, V , and its relative velocity, v , with respect to the medium

$$NV^2 = nv^2, \quad (3)$$

or

$$T \sim \frac{m}{k} \frac{n}{N} v^2 \sim 6 \times 10^{-9} \frac{n}{N} v^2, \quad (4)$$

if v is in cm./sec.* The radiation by the shell has been neglected.

At first the temperature and density distribution in the shell will be very complicated, but it seems probable that, after a certain time, some kind of semi-stationary state may be set up, in which the shell's temperature is approximately that indicated by (3), while its density will fall off exponentially from the front of the shell inwards. The shell is, so to say, continually being compressed towards its front surface, which itself is being held back by the outside pressure due to its motion through the interstellar medium. If we neglect variations in temperature caused by radiation from the shell, the density distribution in the shell will follow the same law as in the Earth's atmosphere, viz.

$$N(x) = N(0)e^{-\frac{m}{kT}\ddot{X}x}. \quad (5)$$

The x -axis points to the star, $x=0$ is the shell's outer surface, N is again the density, while $\ddot{X}$ is the shell's deceleration, counted as a positive number.* If we define the effective thickness of the shell, d , as the value of x at which the density has dropped to $1/e$ times its value at the "outer surface", we find from the above formula that

$$d = \frac{kT}{m\ddot{X}}. \quad (6)$$

The excitation mechanisms in a shell excited by its motion through the interstellar gas differ to some extent from those in a shell excited by the radiation of a star, which has heretofore been exclusively considered. In particular, the direct collisional excitation of the various levels of the hydrogen and helium atoms will now play an important part. For the lower levels they will be more important than the recombinations, and the intensity of H_α will therefore be enhanced as compared to the higher members of the Balmer series. Limitations of time forbid to enter upon details of the intensity computations. But I wish to draw attention to another difference between the spectrum of a shell excited by stellar radiation and one excited by the interstellar gas. It has been shown by several investigators, and in particular by the extensive investigations by Menzel and his co-workers, that the electron temperature of a shell excited by radiation, like a planetary nebula, will not in general become higher than about 9000° , as a consequence of the frequent inelastic encounters of electrons with oxygen ions. In a shell excited by interstellar friction, however, there is no such limitation for the electron temperature, the electrons being practically kept at the temperature given by (4) regardless of the encounters with oxygen ions. We may thus expect to find frequently in these latter shells considerably higher electron temperatures. Higher electron temperatures will reveal themselves by a strengthening of the forbidden lines with respect to the recombination lines, that is, by an intensification of all other lines relative to those of hydrogen and helium. This strengthening may be very pronounced, and may thus aid in distinguishing, at least in some cases, the two types of excitation.

* Since the text for this lecture was written Professor Burgers has derived more complete formulæ. See *Proc. Akad. Sci. Amst.*, **49**, 589, 1946.

Let us now try to make some numerical estimates for cases that may actually present themselves. I propose to consider in particular the shells around Nova Persei 1901 and the Cygnus loop. It may be remarked in passing that the interstellar medium cannot, in general, have any effect on the shells of planetary nebulae.

The mass of an average nova shell has been estimated by various authors to be of the order of 10^{-5} that of the Sun.* The theory underlying these estimates is somewhat uncertain, but their general order is probably correct. The masses of individual shells are likely to show a considerable range. Gordeladse estimates, for instance, that for Nova Persei 1901 the mass would be $3 \times 10^{-6} \odot$, and for Nova Aquilæ 1918 $1.0 \times 10^{-4} \odot$.

If expanding in a region of average interstellar density these nova shells rapidly decelerate, as is indicated in Table IV. Under t this shows the time interval after which the shell has been decelerated to half, and one hundredth, respectively, of its original velocity; for the latter a standard value of 1000 km./sec. was used. The computations have been made with values of $10^{-5} \odot$ and $10^{-4} \odot$ for the shell's mass, as well as for a mass of $10 \odot$, in order to represent the supernova shells. The interstellar density has been taken 3×10^{-24} . The table also shows the radii, R , and effective thicknesses, d , of the shells. The latter are proportional to the temperature, the values in the table are for a temperature of $10,000^\circ$. All data are nothing but rough estimates from a very much simplified model.

TABLE IV

Deceleration of nova shells

	Original mass $10^{-5} \odot$		Original mass $10^{-4} \odot$		Original mass $10 \odot$	
	$v/v_0 = \frac{1}{2}$	$v/v_0 = 1/100$	$v/v_0 = \frac{1}{2}$	$v/v_0 = 1/100$	$v/v_0 = \frac{1}{2}$	$v/v_0 = 1/100$
t (years)	48	4500	102	9800	4800	450 000
R (parsecs)	0.04	0.18	0.08	0.38	3.8	17.6
d (parsecs)	0.000 01	0.05	0.000 02	0.11	0.000 8	4.9
d/R	0.000 2	0.3	0.000 2	0.3	0.000 2	0.3

We see from the table that, in case a nova shell expands in an interstellar medium of average density the deceleration should generally become observable within half a century. After five or ten thousand years its velocity will have decreased so much that it will dissipate. For supernovæ the deceleration may be expected to become visible only after several thousands of years.

It is by no means certain that deceleration effects of this order will actually be observed, for it may be that the average interstellar density in between the cloud systems is considerably less than the average; we have no information on this. It might also be supposed that the interstellar gas has been driven away from the vicinity of a nova by radiation pressure or by previous shells. The former seems improbable; as to the latter, we have no sufficient information regarding the period between successive outbursts. As regards general order a shell might just about have come to rest with respect to the interstellar medium and be overtaken by the nova, in the interval between two outbursts.

Observations of nebulous shells extending over intervals sufficiently long to be

* Cf. Ambarzumian and Kosirev, *Z. Astrophys.*, **7**, 320, 1933; Gordeladse, *Bull. Astrophys. Obs., Abastumani*, **1**, 75, 1937; Grottrian, *Z. Astrophys.*, **13**, 215, 1937 (Nova Herculis 1934).

of interest in the present connection are available for two ordinary novæ: Nova Persei 1901 and Nova Aquilæ 1918. The regular ring-shaped nebula around the latter has shown no signs of deceleration according to recent information from Baade. If we take Gordeladse's mass estimate of $10^{-4} \odot$, this is in agreement with expectation.

The shell around Nova Persei 1901 has been observed considerably longer. This circumstance, coupled with the fact that the nova is evidently in a region with fairly dense interstellar clouds, gives a certain a priori probability that strong deceleration effects might be observed. The evidence available is shown in Plate 6B, giving respectively Ritchey's plate of 1917 November taken with the 60-inch telescope, a plate taken by Hubble with the 100-inch telescope in 1934 October, and one taken with the same instrument by Baade in 1943 September. I am much indebted to Hubble and Baade for their permission to show these latter plates. The three plates have been reduced to the same scale. In the lower left-hand corner a reproduction has been given, on a scale 14 times smaller, of the drawing made by Ritchey after his plate of 1901 November 13. In the 1917 picture we observe an irregular ring structure, and, situated outside this ring, a much brighter line segment. In the later pictures the ring is no longer visible, but the line segment has remained the most striking feature. We have tried to draw rough isophotes in the 1917 picture. If the expansion had proceeded at a uniform rate, dimensions would have increased twice by 1934. The isophotes corresponding to this supposition have been drawn in the 1934 picture. It looks very much as though the bright bar has been strongly braked by the interstellar gas. Though the appearance seems quite strong, I am somewhat hesitant in putting forward this suggestion; it must be admitted that it is conceivable that the deceleration is only apparent, and that in reality we observe different material in the two cases. Baade himself preferred this latter interpretation, especially because in later years, as shown by the 1943 picture, other material did become visible. Yet, I find it hard to adopt this point of view. There is too good an agreement between the position angles of the ends of the bar; and how could we explain, on the hypothesis of uniform expansion, that the bar shown in Hubble's picture of 1934 is not prominent in Ritchey's 1917 plate. And why would Ritchey's bar have completely ceased to be visible in 1934? Lastly, we should point out that the strong deceleration indicated by these two pictures is quite of the order of what was to be expected, with Gordeladse's mass estimate and an evidently above-normal interstellar density.

In my opinion the objections mentioned are serious enough, but there are two more grave objections to the classical picture of a uniformly expanding nebula excited by the old nova. The gravest is the existence of the bar itself. It is hard to conceive how a nova would at the same time have ejected the more or less symmetrical shell shown on the 1917 plate, and this most one-sided bar. This latter is certainly quite unusual; nothing like it has been reported for any of the other five novæ around which shells have been photographed. The second objection is founded on Baade's 1943 photograph, which shows newly luminous matter in position angles where little or nothing could be seen on the 1934 plates, and on Baade's observation "that the integrated brightness of the shell is today much higher than it was in 1917".*

* This was communicated to me in a letter from Baade.

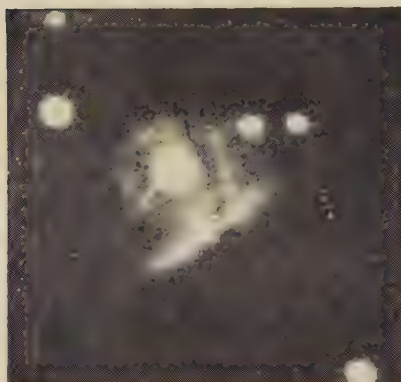


Nebulosity near ζ Orionis (100-inch Mount Wilson).

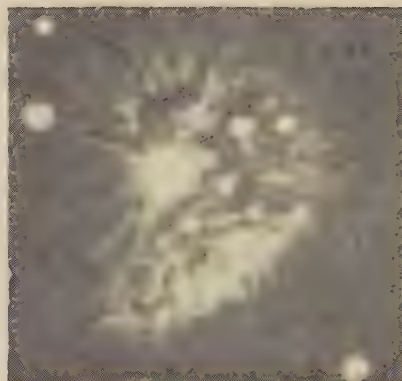
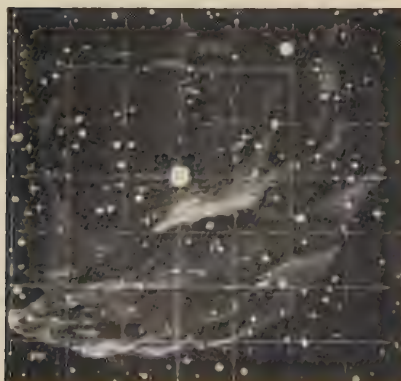
A



NGC 4594. 60-inch Mount Wilson.



B



Nova Persei 1901.

Upper left: Ritchey, 1917 Nov. 15, 60-inch Mount Wilson.

Upper right: Hubble, 1934 Oct. 3, 100-inch " "

Lower right: Baade, 1943 Sept. 28, 100-inch " "

Lower left: Ritchey, 1901 Nov. 13, drawing, on 14 times smaller scale than the 3 preceding reproductions.

All these facts can be readily understood if we adopt as working hypothesis that the shells thrown out at the outburst of Nova Persei 1901 were more or less symmetrical. The main shell, that corresponding to the so-called principal spectrum, is shown as the fine ring in Ritchey's picture, while at the later epochs it has become too faint to be visible. This shell was made luminous in the usual manner by the radiation from the old nova. It is known from several investigations, and in particular from those by McLaughlin*, that novæ also throw off a shell with a velocity of about twice that of the main shell. These shells, which appear to be less dense, are represented a little while after the outburst by the so-called "diffuse enhanced spectrum", which has been identified also in Nova Persei 1901. I wish to propose the hypothesis that the bar is that part of the outer shell that has run up against dense interstellar matter, that it derives its luminosity from the resulting rise in temperature, and that, at the same time, it is braked by the interstellar matter. It is to be noted, in this connection, that the interstellar matter is arranged curiously asymmetrically with respect to the nova, extending in a fan-like manner up to the star just in that quadrant where the bar is situated. The increase in brightness noted by Baade is not unexpected for a shell excited by the interstellar matter. Furthermore, it is easy to see how new material could become visible, as other parts of the shells run up against the interstellar medium.

I cannot now enter into details, nor into the interesting question whether the curious arrangement of the interstellar matter in the nova's immediate vicinity might be explained as due to previous nova shells. I only wish to point out that, if the shell has been decelerated, Humason's computations of the distance of the nova must be revised; the distance would become about 900 parsecs instead of 600; the absolute magnitude at maximum, -9.7 , would be still more exceptional for an ordinary nova.

The truth of the hypothesis should be tested in two ways. From the deceleration we can roughly determine the amount of internal energy which the shell gains each second. This should have to be at least equal to the amount radiated. The still rough calculations made so far show that the two amounts are of the same order. Also, the computed amount of radiation in the photographic region does not seem to disagree with the estimated photographic magnitude of the nebula.

The second test concerns the spectrum. If the collisional mechanism is responsible for the radiation, H_{α} should be considerably enhanced as compared with the higher Balmer lines. I have no numerical observational information on this point, Humason's intensity estimates† extending only to $\lambda 5000$. But the spectrum, according to Humason, shows another remarkable anomaly, namely that the lines of both hydrogen and helium are extremely weak relative to the forbidden lines. A comparison with the intensity estimates in planetary nebulae shows that, for equal intensities of the forbidden lines in the two cases, the permitted hydrogen and helium lines in the Nova Persei nebula are some 15 times fainter than in planetary nebulae. If the Nova Persei nebula in 1934 were excited by radiation from the old nova this would mean that the relative abundance of both hydrogen and helium as compared with oxygen and neon is 15 times lower than normal. Though this possibility cannot be excluded, it is rendered improbable

* Cf. *Ap. J.*, **95**, 428, 1942.

† *Ap. J.*, **88**, 228, 1938; (*Contr. Mt. Wilson Obs.*, No. 596.)

by the fact that hydrogen and helium lines seem to have been normally prominent in the nova's spectrum after the outburst. A more satisfactory explanation is offered by the hypothesis that the nebula is excited by its motion through the interstellar medium. In this case the electron temperature, as indicated by the level of ionization, would be at least $30,000^{\circ}$. At that temperature the forbidden lines would be enhanced relative to the hydrogen and helium lines in just about the ratio observed by Humason.

There are two other cases of moving shells which are of interest in the present connection, namely the Veil or Network nebulae in Cygnus, and the Crab nebula.

The striking similarity in structure of the two principal parts of the Veil nebulae, their curved arrangement, and particularly the curvature of most of their separate wisps, can leave no reasonable doubt that the nebulosities in this region are parts of a shell expelled from a star.* Hubble has observed that the shell is still expanding, the velocity of expansion being about $0''.03$ per year. In measuring faint wisps of nebulosity inside the loop Humason† has found indications of a linear velocity of expansion of the order of 75 km./sec. This value seems still very uncertain; if it is of the right order the distance would be about 500 parsecs, and the radius of the shell about 10 parsecs. Searches down to $12^m.5$ have failed to show a central star which could be responsible for the bright line radiation of these nebulae. NGC 6960 is situated at the border of a dark nebula; it is unknown whether there is any real connection between the two.

The feature of the Veil nebulae which is the most puzzling, when one comes to think of it, is the extreme thinness of many of their wisps, which does not surpass 2 or 3" in some cases. If the nebulae were freely moving through space it would not seem understandable how parts of them could have remained anything like as thin during the time the shell has needed to expand to its present size. The thinness of the wisps would appear to point very definitely to a compression due to their motion through the interstellar medium. Rough estimates by Burgers indicate that the observed values may not be incompatible with theory.

It is plausible then, also to look for the source of excitation in the friction with the interstellar gas. If the interstellar density is not too far below the average this could indeed raise the temperature sufficiently to explain the observed radiation. I cannot now enter upon the details of the discussion of this case. The mass of the whole shell must be quite considerable, of the order of that of the Sun, thus more resembling that of a planetary nebula or of a supernova shell (like the Crab nebula) than that of an ordinary nova. Even apart from their exceptional size the Cygnus nebulae do not, however, resemble an ordinary planetary, because of the thinness of the shell, its wispy structure, and the faintness of the central star. A supernova shell seems more probable; the originally large velocity of expansion would then have been braked down to the present value by the interstellar medium. I found that this suggestion had already been made by Zwicky.‡

In conclusion a few words about the Crab nebula, of which Baade and

* For pictures of these nebulae the reader may be referred to *Publ. Lick Obs.*, **11**, Pl. 80, or to Ross's *Atlas of the Milky Way*, Pl. 17. Details are shown in the beautiful pictures by Duncan, *Ap. J.*, **57**, 146, 1923; (*Contr. Mt. Wilson Obs.*, No. 256), and *Ap. J.*, **86**, 496, 1937; (*Contr. Mt. Wilson Obs.*, No. 579).

† Private communication.

‡ *Rev. mod. Phys.*, **12**, 72, 1940.

Minkowski have recently made such interesting studies.* The mass is so large in this case that no deceleration can be expected, and indeed none has been observed. If the interstellar density were about normal the nebula might have been heated to the temperature of the order of a few hundred thousand degrees which would be required to explain why the inner amorphous nebulosity shows no hydrogen and helium lines. The hypothesis that also in this case the density distribution in the shell and the radiation are determined by the interaction with the interstellar medium is attractive, primarily because it might explain how the exploded star can have been transformed into a shell of such surprising homogeneity as the main amorphous mass of the Crab nebula. However, the detached filamentary outer structure, whose relation to the inner mass has been so beautifully revealed by Baade's photographs in various wave-lengths, seems to put serious difficulties in the way of this hypothesis.

It has been my purpose to expose to you this afternoon some phenomena connected with interstellar matter. I am afraid I have not been able more than to sketch a *possible* way in which some of the problems that offer themselves might be solved.

* *Ap. J.*, 96, 188 and 199, 1942 ; (*Contr. Mt. Wilson Obs.*, Nos. 665 and 666).

ON THE SPIRAL CHARACTER OF THE EXTERNAL GALAXIES

E. A. Milne, F.R.S.

(Received 1946 February 26)

Summary

The orbits of free particles in the neighbourhood of a massive nebular nucleus are calculated on the kinematic theory of gravitation. They are shown to be Keplerian orbits on the dynamic time-scale, but open equi-angular spirals on the absolute or kinematic time-scale, and it is shown how they would be so recognized by an observer at the nucleus. The suggestion of E. W. Brown, that nebular arms are the envelopes of actual orbits, is examined in detail, and shown to be untenable. It is shown in fact that the actual case must be the reverse, namely that the orbits *cannot* have an envelope. The condition for this is found, and shown to give the present positions of the particles forming a nebular arm. This locus satisfies the functional equation for a nebular arm, and corresponds to ejection from a fixed point near the centre of the nebula, in accordance with Jeans's ideas and with a conclusion of Lindblad's. The theoretical locus unwinds for a finite number of convolutions in the sense of the orbital motion, and then doubles back, unwinding thereafter for the same number of convolutions in the retrograde sense. It is accordingly suggested that the spiral structure consists of two parts, one interior to a critical distance, the other exterior. It is further suggested that "barred" spirals are systems for which ν_0 , the present number of complete convolutions, is small, and for which much of the material lies near the critical distance. Comparison is made with observation.

1. According to Hubble*, 80 per cent. of the external galaxies are classified as spirals, of one type or another. Yet it is remarkable that at present there is no generally accepted explanation of the spiral character of the external galaxies. With the coming into operation of the 200-inch telescope on Mt. Palomar, it is to be expected that more detailed knowledge will become available of the true structure of these objects, and the advantage of a working model needs no stressing. The present paper endeavours to provide such a working model, based essentially on the classical theory of gravitation, but modified so as to take into account the consequences of my recent kinematic studies in the nature of gravitation. The paper is essentially a study of the orbits of particles in the presence of a massive nucleus, but it is realized that this idealization does not do justice to the observed disk-like configurations of galaxies. Accordingly no finality is claimed for the model. Nevertheless a study of the orbits in the presence of a massive nucleus is an indispensable preliminary to the undertaking of a similar study for a nuclear structure more approaching reality. It is believed accordingly that the present attempt is worth while, and that it gives insight into the observed structure. It is recognized that the attempt involves a considerable degree of abstraction and simplification, but the essence of pioneering attacks is to introduce the right amount of simplification.

* Hubble, *The Realm of the Nebulae*, p. 55, Table IV, 1936. He is quoting from the Harvard Survey of bright nebulae covering the entire sky.

2. The spiral character exhibited by so many of the external galaxies has been an object of speculation from the time of its discovery by Lord Rosse (1850). A series of papers by Stephen Alexander* in 1852 attempted to link it up with the forms of rotating liquids following disruption due to instability. Wilczynski† in 1896 attempted to connect the spiral form with the varying angular velocity in a set of Keplerian orbits round a nucleus, pointing out that a set of particles initially distributed along a radius will be bent into an arc of a spiral in the course of successive revolutions. A series of papers by Lindblad‡ has considered in detail the stability of rotating systems composed of stars, especially the effects of a disturbance on motions on the verge of instability in an equatorial rim. He has claimed thus to account for the spiral forms of arms, but I am not aware that he has constructed a quantitative model of a spiral nebula. The notable researches of Jeans§, summed up in his two comprehensive volumes, are fresh in all our memories. It was his suggestion that the occurrence of spiral arms as a pair emerging from opposite ends of a nuclear diameter may be due to tidal forces arising from the gravitational effects of the rest of the universe. Jeans, however, was unable to explain the actual spiral character, and was indeed sceptical of the possibility of explaining it within the confines of Newtonian physics. He summed up his views on the state of knowledge in 1928 in the following words||:

“The only result that seems to emerge with some clearness is that the spiral arms are permanent features of the nebulae Their further interpretation forms one of the most puzzling, as well as disconcerting, problems of cosmogony Each failure to explain the spiral arms makes it more and more difficult to resist a suspicion that the spiral nebulae are the seat of types of forces entirely unknown to us, forces which may possibly express novel and unsuspected metric properties of space. The type of conjecture which presents itself, somewhat insistently, is that the centres of the nebulae are of the nature of singular points at which matter is poured into our universe from some other, and entirely extraneous, spatial dimension, so that, to a denizen of our universe they appear as points at which matter is continually being created”.

3. The conclusions to which I have been led by my studies in cosmology are akin to these speculations of Jeans. Instead of a singular point at the centre of each nebula, we have a singular event, at the natural origin of time, an event which may equally well be supposed to have taken place at the centre of any one nebula, the various nebulae subsequently separating by their mutual recession.¶ At this singular epoch, the creation of the whole system took place. The spiral form then ultimately arises from the unsuspected new properties of gravitation which have emerged from my kinematic studies, chief amongst which is the secular dependence of the so-called “constant” of gravitation on the absolute epoch.

* *A. J.*, 2, 1852.

† *Ap. J.*, 4, 97, 1896.

‡ *Stockh. Obs. Ann.*, *passim*. In particular 12, no. 4, 1936, and 13, no. 5, 1940 (esp. pp. 30 ff.). See also *Ap. J.*, 92, 1, 1940.

§ *Problems of Cosmogony and Stellar Dynamics*, Cambridge, 1919; *Astronomy and Cosmogony* Cambridge, 1928.

|| *Astronomy and Cosmogony*, p. 352.

¶ *Relativity, Gravitation and World-Structure*, p. 185, Chap. X, Oxford, 1935.

4. I have previously considered the question of the spiral form of the external galaxies in my paper, "The Inverse Square Law of Gravitation, II"*[†], wherein I showed that orbits, considered as circular on the ephemeral τ -scale of time, would appear as equiangular spirals on the absolute or kinematic t -scale. I wrote: "Actually we do not know whether the spiral coils of nebulae represent orbits or not. We only know that we are seeing a set of particles or stars strung out along their orbits. But our analysis seems to indicate the origin of the spiral character of resolved galaxies". I then went on to discuss some of the difficulties associated with the dual description of orbits as apparently closed (on the τ -scale) and open, on the t -scale, and concluded: "Clearly great care will be required to disentangle all these effects. We may be content at this stage with our general tentative explanation of the spiral character of resolved nebulae as connected with the fact that the orbits being described are spirals when reckoned in t -measure, which is the measure in terms of which the nebulae are receding The elucidation of further details can best be dealt with in an astronomical paper".

The present paper is the paper referred to. Since 1937 I have not returned to the subject till now, save that in my first Presidential Address to the Society, "On the Nature of Universal Gravitation", I wrote[‡]: "I have throughout referred to the galaxies as particles. But the $t \rightarrow \tau$ transformation is to be expected to throw great light on their spiral structure, inasmuch as gravitational orbits, regarded dynamically as *closed*, will be regarded as open on the t -scale". My return to the subject I owe directly to Dr Martin C. Johnson's recent volume[§], in Appendix II of which he discusses the application of my theory of gravitation to spiral nebulae in some detail. Johnson's conclusions have been criticized by A. G. Walker[¶], whose difficulties I respect. The present paper will be found, I hope, to constitute a sufficient answer to Dr Walker.

5. In the present paper I do not assume that the spiral arms indicate actual orbits; instead, I use them for what they are—the loci of the present || positions of the particles or stars composing them. I shall derive an equation for the curve shown by the spiral arms, by three different methods: the first was suggested by some analysis due to E. W. Brown[¶]; the second depends on the character of *permanency* possessed by the arms (stressed by Jeans in the above quotation); the third expresses in analytic form the consequences of emission of particles, during past time, from a fixed point in the nebula. All three methods yield the same equation for an arm. This equation determines the order of magnitude of the size of a nebula, the past epoch at which emission began, and the number of convolutions which should have been formed since then.

Throughout, I use my own kinematic theory of gravitation. But the theory which follows, though arising out of kinematic relativity, is in no way an essential consequence of kinematic relativity, since it is in part empirical, and will probably need revision in the light of future observations.

6. I begin with the fundamental idealized problem, namely the problem of the orbits described by particles in the vicinity of a massive nucleus. The

* *Proc. Roy. Soc.*, **160**, A, 18, 1937. Referred to as *Gravitation, II*.

† *M.N.*, **104**, 134, 1944.

‡ *Time, Knowledge and the Nebulae*, pp. 161–183, 1945.

§ *Observatory*, **66**, 215, 1946. Dr Walker courteously sent me a copy of his "letter" before publication.

|| Apart from a correction depending on the time of travel of light from them to us, which is irrelevant in the present investigation.

¶ *Ap. J.*, **61**, 111, 1925, "Gravitational forces in spiral nebulae".

derivation of the t -equations of the actual orbits has been given elsewhere, but in a publication not always accessible in observatories, and I accordingly re-sketch it here. Consider a nebular nucleus of mass M , in the presence of which, and in the presence of the rest of the universe, a free particle P is describing an orbit. This nebular nucleus, O , is taken to be a fundamental particle, and is taken as origin. The 4-vector equations of motion of P have been shown* to be

$$\frac{\xi^{\frac{1}{2}}}{Y^{\frac{1}{2}}} \frac{d}{dt} \left(\frac{V}{Y^{\frac{1}{2}}} \right) = - \frac{\xi^{\frac{1}{2}}}{X} \left(\mathbf{P} - V \frac{Z}{Y} \right) - \frac{Mc^3 t}{M_0} \frac{\mathbf{P}}{|\mathbf{P}|^3} + \frac{V}{Y^{\frac{1}{2}}} \frac{1}{Y^{\frac{1}{2}}} \frac{d}{dt} (\xi^{\frac{1}{2}}), \quad (1)$$

$$\frac{\xi^{\frac{1}{2}}}{Y^{\frac{1}{2}}} \frac{d}{dt} \left(\frac{c}{Y^{\frac{1}{2}}} \right) = - \frac{\xi^{\frac{1}{2}}}{X} \left(ct - c \frac{Z}{Y} \right) - \frac{Mc^2}{M_0} \frac{1}{|\mathbf{P}|} + \frac{c}{Y^{\frac{1}{2}}} \frac{1}{Y^{\frac{1}{2}}} \frac{d}{dt} (\xi^{\frac{1}{2}}), \quad (2)$$

where $\mathbf{P}$, $\mathbf{V}$ are the position-vector and velocity-vector of the particle P with regard to O , t is the absolute value of the epoch considered, reckoned from the absolute zero of time, M_0 is the apparent mass of the smoothed-out universe, and

$$X = t^2 - \mathbf{P}^2/c^2, \quad Y = 1 - \mathbf{V}^2/c^2, \quad Z = t - \mathbf{P} \cdot \mathbf{V}/c^2, \quad \xi = Z^2/XY.$$

Equation (2) is deducible from equation (1) on scalar multiplication by $\mathbf{P}/ct$ and some consequential simplification, but the full power of the 4-vector analysis is only realized by employing both (1) and (2); (2) is of course the time-component corresponding to the 3-vector equation (1). A 3-vector equation of interest on its own merits follows by multiplying (2) by $\mathbf{V}/c$ and subtracting from (1), when we get

$$\frac{d\mathbf{V}}{dt} = - \frac{Y}{X} (\mathbf{P} - \mathbf{V}t) - \frac{Mc^3 t}{M_0 \xi^{\frac{1}{2}}} \frac{Y}{|\mathbf{P}|^2} \left[\frac{\mathbf{P}}{|\mathbf{P}|} - \frac{\mathbf{V}}{c} \frac{|\mathbf{P}|}{ct} \right]. \quad (3)$$

In (1), the first term on the right represents the pull of the smoothed-out universe on the eccentrically-situated particle P , the second term on the right represents the inverse-square attraction due to the nearby nucleus M at O , and the third term on the right represents certain effects due to change of mass with velocity. The form of (3) shows that when P is itself a fundamental particle, satisfying $\mathbf{P} = \mathbf{V}t$, the pull of the smoothed-out universe vanishes, and, moreover, that the inverse-square attraction term vanishes when P is on the confines of the expanding universe, so that $t^2 = \mathbf{P}^2/c^2$.

By vectorial multiplication of (1) by $\mathbf{P}$, it can be shown to admit the exact integral

$$\mathbf{P} \wedge \mathbf{V} = \text{const.} \times Z. \quad (4)^\dagger$$

It is more convenient in what follows to use polar co-ordinates (r, θ) , in the plane of the orbits, where $r = |\mathbf{P}|$. On neglecting $|\mathbf{V}|$ compared with c and $|\mathbf{P}|$ compared with ct , (3) takes the form, as regards its radial component,

$$\ddot{r} - r\dot{\theta}^2 = - \frac{r - t\dot{r}}{t^2} - \frac{c^3 t}{M_0} \frac{M}{r^2}, \quad (5)$$

* *Gravitation II*, equations (54), (54').

† It has often been stated in criticism of my cosmology that it applies only to a homogeneous universe. Actually it applies equally well to a universe containing any density-irregularities whatever: we merely need to superpose on the general field of acceleration produced by the averaged-out homogeneous distribution of matter the local gravitational effects of each separate nebular nucleus, by formulæ of the type of (3). The reason is that what at an embryonic stage of the theory were interpreted as uniform distributions of *matter* are seen in retrospect to be essentially uniformly distributed *frames of reference*. The theory developed in fact more rapidly than the criticisms.

whilst the integral (4) reduces under the same circumstances to

$$r^2 \dot{\theta} = \text{const.} \times t. \quad (6)$$

7. Let h_0 be the angular momentum of P per unit mass at epoch t_0 . Then (6) gives

$$r^2 \frac{d\theta}{dt} = h_0 \frac{t}{t_0}. \quad (7)$$

To solve (5), make the substitution from r, t to ρ, τ given by

$$r = \rho \frac{t}{t_0}, \quad \frac{dt}{t} = \frac{d\tau}{t_0}. \quad (8), (9)$$

In (7), (8) and (9), t_0 denotes an arbitrarily given fixed epoch, which we shall usually take to be the *present* epoch. The integral of (9) we shall write in the form

$$\tau = t_0 \log(t/t_0) + t_0, \quad (10)$$

so that at the epoch $t = t_0$, we have $\tau = t_0$ and $d\tau = dt$.

By direct differentiation of (8) we have

$$\dot{r} = \frac{1}{t_0} \left[\rho + t \frac{d\rho}{dt} \right] = \frac{r}{t} + \frac{d\rho}{d\tau},$$

so that

$$-\frac{r - t\dot{r}}{t^2} = \frac{1}{t} \frac{d\rho}{d\tau}.$$

Further

$$\ddot{r} = -\frac{r}{t^2} + \frac{1}{t} \left[\frac{r}{t} + \frac{d\rho}{d\tau} \right] + \frac{t_0}{t} \frac{d^2\rho}{d\tau^2} = \frac{1}{t} \left[\frac{d\rho}{d\tau} + t_0 \frac{d^2\rho}{d\tau^2} \right].$$

Hence the radial equation of motion (5) yields

$$\frac{t_0}{t} \frac{d^2\rho}{d\tau^2} - \rho \left(\frac{t}{t_0} \right) \left(\frac{d\theta}{d\tau} \right)^2 \left(\frac{t_0}{t} \right)^2 = -\frac{c^3 t}{M_0} \frac{M}{\rho^2} \left(\frac{t_0}{t} \right)^2,$$

or, dividing through by the factor t_0/t ,

$$\frac{d^2\rho}{d\tau^2} - \rho \left(\frac{d\theta}{d\tau} \right)^2 = -\frac{c^3 t_0}{M_0} \frac{M}{\rho^2}. \quad (11)$$

Put

$$\gamma_t = c^3 t / M_0, \quad (12)$$

$$\gamma_0 = c^3 t_0 / M_0. \quad (13)$$

Then (11) becomes

$$\frac{d^2\rho}{d\tau^2} - \rho \left(\frac{d\theta}{d\tau} \right)^2 = -\gamma_0 \frac{M}{\rho^2}. \quad (14)$$

The integral (7) yields similarly

$$\left(\rho \frac{t}{t_0} \right)^2 \frac{d\theta}{d\tau} \left(\frac{t_0}{t} \right) = h_0 \frac{t}{t_0},$$

or, dividing by the factor t/t_0 ,

$$\rho^2 \frac{d\theta}{d\tau} = h_0. \quad (15)$$

8. We see that the transformation $t \rightarrow \tau$, with the associated transformation $r \rightarrow \rho$, removes the cosmical term from the right-hand side of (5), transforms the coefficient γ_i of the inverse-square term (representing the nuclear attraction) into a constant γ_0 , and transforms the angular momentum $h_0(t/t_0)$ into a constant h_0 . We recognize (14) and (15) as the Keplerian equations for motion about a nucleus of mass M under the classical Newtonian law of gravitation, with constant of gravitation γ_0 and with angular momentum h_0 per unit of mass.

9. The success of this transformation is of course independent of my well-known interpretation of the formulæ (8), (9) and (10) in terms of light-signals and re-graduated clocks. Nevertheless its only rational meaning is that if t is the time-variable used in Lorentz formulæ and Lorentz invariant formulæ such as those from which (1) and (2) were derived, then τ must be the time-variable of current dynamics*; and further that if the standard metre measures ρ , as it must do for (14) and (15) to be the Keplerian equations of motion, then by (8), in a given distance r on the t -scale, the value attached to ρ steadily decreases as t advances, and so in the t -scale the standard metre is expanding. Conversely, a constant length ρ as measured by the standard metre implies an increasing measure r on the t -scale.

It is now necessary to emphasize that there is not just *one* τ -scale, but a continuum of τ -scales, depending on choice of the constant t_0 . We should really write the transformation not just $t \rightarrow \tau$ but $t \rightarrow \tau(t_0)$. We usually make the t - and τ -scales agree at the present moment $t = t_0$, but we can make them agree at any chosen epoch t_0 . The $\tau(t_0)$ scale is then tangent to the t -scale at $t = t_0$. This is depicted in the diagram.

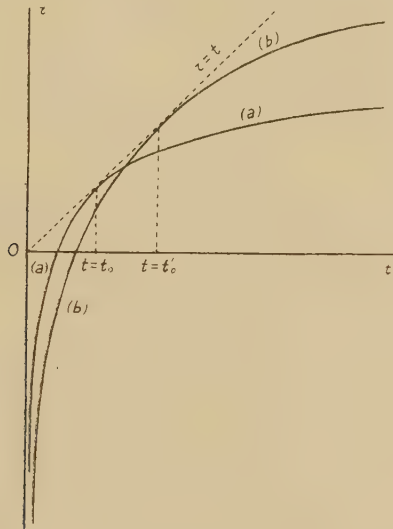


FIG. 1.—Time-scales for different values of the parameter t_0 .

- (a) $\tau = t_0 \log (t/t_0) + t_0$.
- (b) $\tau = t'_0 \log (t/t'_0) + t'_0$.

* I regard the t -scale as more fundamental than the τ -scale, because the various equations, etc., on the t -scale do not involve the parameter t_0 , whilst the τ -equations do. The parameter t_0 cannot be understood without prior reference to the t -scale.

The value attributed to the "constant" of gravitation, namely γ_0 , depends on the value chosen for t_0 . In calculations made by any one observer using a definite $\tau(t_0)$ time-scale, the constant of gravitation will behave as a definite constant, with value γ_0 ; but another observer at another epoch, using a τ -clock appropriate to that epoch, will find another value for γ . Any one τ -scale moves eventually out of coincidence with the t -scale.

Likewise, for any one observer the angular momentum h_0 will behave as if constant. But another observer, at a later epoch t_1 , using the non-appropriate τ -measure of time, will find the new value $h_0(t_1/t_0)$ for the angular momentum. It will again behave as if constant, but the value attributed to it will be different. It will be seen that in spite of the constancy of angular momentum for calculation purposes, in calculations by any one observer using any one τ -scale, the angular momentum h , like γ , *really* increases with the epoch. Thus the present analysis accounts of itself for the prevalence of *spin* in the universe, its secular increase arising from the pull of the rest of the universe on a particle eccentrically situated with respect to a fundamental particle.

10. A more striking illustration of the effect of the $t \rightarrow \tau(t_0)$ transformation is given by consideration of the *distance* of a nebular nucleus from an observer situated on another nebular nucleus. Let the first nucleus have a recession-velocity V_0 . Then on the t -scale its distance is the increasing distance $r = V_0 t$, at epoch t . When the observer passes to the $\tau(t_0)$ scale, he attributes to the distant nucleus the value $\rho = (V_0 t)(t_0/t) = V_0 t_0$, which is constant*; but different observers at different stages of their experience will attribute different distances $V_0 t_0$ to it, because they will use different values for t_0 .

Thus though the $t \rightarrow \tau(t_0)$ transformation destroys the relative motion, a τ -observer will be able, over long periods of time, to detect the motion, because he will attribute different constant distances to the distant nucleus from time to time. With any one clock, graduated to τ , he will find a perfectly consistent Newtonian dynamics and gravitation; but as time goes on, he will be disposed to re-adjust his clock-graduations to some new value of t_0 , and alter his attributions of distance accordingly.

In the past it has been necessary to stress the *constancy* of γ_0 , h_0 , $V_0 t_0$ for an observer using $\tau(t_0)$ time, because for all purposes of practical physics we use a fixed value of t_0 , namely the present epoch on the t -scale. But the $\tau(t_0)$ scale is an ephemeral scale, in contrast to the absolute or t -scale, and when, as in the theory of the spiral nebulae, it is necessary to consider phenomena over long periods of time, it becomes necessary to stress the secular changes in the various constants.

11. Consider now the implications of the transformation for the appearance of orbits in spiral nebulae to an observer. Consider first an observer in the plane of the galaxy in question. The solutions of (14) and (15) are well known to be Keplerian ellipses. That means that our observer, armed with the usual expanding metre of ordinary physics, and measuring the orbit of a particle continually as time progresses, will find it a re-entrant ellipse. But by the time that the particle has completed a revolution, the metre-scale will be bigger than it was at the beginning, and so the orbit is actually non-re-entrant.

* The exact formula is $\frac{1}{2} c t_0 \log \left(\frac{1 + V_0/c}{1 - V_0/c} \right)$.

Next, consider the appearance to another observer on a distant nucleus. At epoch t , he will consider the angle ϵ subtended by the radius-vector in the plane of the nucleus (supposed for simplicity to be perpendicular to the line of sight) to be given by $\epsilon = r/V_0 t$; using $\tau(t_0)$ measure, he will calculate it as $\epsilon_0 = \rho/V_0 t_0$. By (8), $\epsilon = \epsilon_0$, as pointed out by A. G. Walker. Hence the distant observer will view the orbit as re-entrant, as projected on the sky. But when using t -measure, he will still regard the orbit as an open three-dimensional spiral, now lying on the surface of a cone, the plane through the now receding nucleus intersecting the cone at any moment in a locus containing the particle.

A plane galaxy remains a plane galaxy under transformation to a distant nucleus. For the plane $x = V_0 t$, under the Lorentz transformation to a new origin moving with speed U , namely

$$x = \frac{x' + Ut'}{(1 - U^2/c^2)^{1/2}}, \quad t = \frac{t' + Ux'/c^2}{(1 - U^2/c^2)^{1/2}},$$

gives

$$x' + Ut' = V_0(t' + Ux'/c^2),$$

or

$$x' = \frac{V_0 - U}{1 - UV_0/c^2} t'. \quad (16)$$

This is a particular case of Einstein's velocity-transformation, and it gives the equation of the transformed plane.

The receding plane $x = V_0 t$ intersects the series of opening spiral orbits in a series of points which are the *present* positions of the particles, and it is these we now proceed to discuss. We wish to obtain the present polar co-ordinates (ρ, θ) of the particles constituting a galaxy.

12. It was suggested by E. W. Brown* that a spiral arm of a nebula might be the *envelope* of the orbits of the particles concerned. We shall see shortly that this idea is untenable, but we begin by examining its consequences. Brown pointed out that the polar equation to an orbit contains, in addition to the epoch variable, three parameters, which may conveniently be taken to be a scale-parameter, an orientation-parameter and a shape- or eccentricity-parameter. Now a set of orbits has an envelope only when its equation contains a single parameter. We must therefore impose two relations between the three parameters present. One of these relations may appropriately be taken to be given by putting the eccentricity e equal to zero. It is clear, in fact, that the introduction of a non-zero eccentricity would be equivalent to introducing star-streaming, a phenomenon which is irrelevant in the present survey. To put $e = 0$ is equivalent to restricting attention to *circular* orbits, in τ -measure, as the outstanding systematic motion. The assumption $e = 0$ was in fact made also by Brown.

Once we have taken $e = 0$, the orbits become in τ -measure the circles

$$\rho = \text{const.} = \rho_0. \quad (17)$$

Then (14) gives

$$\rho_0 \left(\frac{d\psi}{d\tau} \right)^2 = \frac{\gamma_0 M}{\rho_0^2}.$$

* *Loc. cit.*

Eliminating $d\theta/d\tau$ between this and (15), we get

$$h_0^2 = \gamma_0 M \rho_0,$$

and then

$$\left(\frac{d\theta}{d\tau}\right) = \left(\frac{\gamma_0 M}{\rho_0^3}\right)^{\frac{1}{2}}. \quad (18)$$

Hence if $\theta = \theta_0$ is the azimuth at epoch t_0 , the motion is given by

$$\theta - \theta_0 = \left(\frac{\gamma_0 M}{\rho_0^3}\right)^{\frac{1}{2}} (\tau - \tau_0).$$

By (10), this may be written

$$\theta - \theta_0 = \left(\frac{\gamma_0 M t_0^2}{\rho_0^3}\right)^{\frac{1}{2}} \log \frac{t}{t_0}, \quad (19)$$

whilst by (8),

$$r = \rho_0 t/t_0. \quad (20)$$

Relations (19) and (20) express the polar co-ordinates of a point on the orbit parametrically in terms of t . Eliminating t , we have the equation to the orbit in the form

$$r = \rho_0 \exp \left[\left(\frac{\rho_0^3}{\gamma_0 M t_0^2} \right)^{\frac{1}{2}} (\theta - \theta_0) \right]. \quad (21)$$

Relation (18) is of course just Kepler's third law for the case of circular orbits, giving in τ -time the period $2\pi(\rho_0^3/\gamma_0 M)^{\frac{1}{2}}$. Relation (21) gives the (r, θ) orbit as an equi-angular spiral, of angle α given by

$$\cot \alpha = \left(\frac{\rho_0^3}{\gamma_0 M t_0^2} \right)^{\frac{1}{2}}. \quad (22)$$

This is the locus to which the circular orbit $\rho = \rho_0$ transforms, once it is recognized that the standard metre is expanding. As t increases, r and θ increase together.

13. In (21) we recognize the presence of two parameters, ρ_0 and θ_0 , ρ_0 being a scale parameter and θ_0 an orientation parameter, in accordance with E. W. Brown's remark. If the orbits (21) are to have an envelope, there must be a relation between ρ_0 and θ_0 —the second of Brown's relations. Such a relation would give, in our notation, the *present positions* of the particles considered.

As an illustrative example, consider the particles which at the present moment are strung out along a radial arm,

$$\theta_0 = \text{const.}$$

These particles will have been describing different orbits, characterized by different values of the parameter ρ_0 , and the envelope of such orbits, if it exists, will be given by eliminating ρ_0 between (21) and the result of differentiating (21) partially with regard to ρ_0 . Differentiating (21) logarithmically, keeping θ_0 constant, we have

$$0 = \frac{1}{\rho_0} + \frac{3}{2} \left(\frac{\rho_0}{\gamma_0 M t_0^2} \right)^{\frac{1}{2}} (\theta - \theta_0). \quad (23)$$

Relations (21) and (23) then give the envelope parametrically in the form

$$\theta = \theta_0 - \frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}}, \quad (24)$$

$$r = \rho_0 e^{-\frac{2}{3}}, \quad (25)$$

ρ_0 being considered as the parameter; and thus, eliminating ρ_0 , the envelope would be

$$\theta = \theta_0 - \frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{r^3 e^2} \right)^{\frac{1}{2}}. \quad (26)$$

This is a spiral type of curve, with r increasing with θ , *i.e.* with the spiral unwinding in the same sense as the orbital motion. The epoch of the point of contact of an orbit with the envelope is given, by (20) and (25), by

$$t = t_0 e^{-\frac{2}{3}} = 0.5134 t_0. \quad (27)$$

But it is now clear that such an envelope could not represent the arm of a spiral nebula. In the first place, it would be entirely artificial to suppose a spiral arm to be a locus touched by orbits along which the *present* positions of the particles are distributed along the radius vector $\theta = \theta_0$. Secondly, the actual epoch of contact would have been in the remote past, at $t = 0.5134 t_0$, where t_0 , the present epoch (age of the universe to us, reckoned from the natural origin of time), is of the order of 2×10^9 years. Thirdly, and most important of all, it would be highly improbable that a set of orbits could pass through their points of contact with their envelope and then reach the line $\theta = \theta_0$ unperturbed; for near the point of contact with the envelope, as recognized by E. W. Brown, the orbits would be exceedingly densely crowded together, and so would subject one another to intense perturbations.

To explore this third argument, let $n_0(\rho_0)d\rho_0$ be the distribution of orbits along the radius vector $\theta = \theta_0$ at the present epoch t_0 , $n(r)dr$ the distribution *in radius vector* at some former epoch t ; then

$$n(r)dr = n_0(\rho_0)d\rho_0.$$

In this relation, dr is the difference in radius vector of the two orbits ρ_0 and $\rho_0 + d\rho_0$ along some radius vector θ , and so

$$n(r) = \frac{n_0(\rho_0)}{(\partial r / \partial \rho_0)}, \quad (28)$$

where $(\partial r / \partial \rho_0)$ denotes differentiation of (21) with respect to ρ_0 , keeping θ and θ_0 both constant. But this denominator *vanishes* at the point of contact of orbit and envelope, and hence the density of orbits in radius-vector tends to infinity near the envelope. It is true that the particles in their orbits do not actually collide—neighbouring particles reach the point of intersection of their orbits at slightly different times—but the orbits are densely crowded together here, and actually touch their envelope all at the same epoch, by (27).

14. The arguments of the last section, though based on the artificial example $\theta_0 = \text{const.}$, are perfectly general. Suppose, quite generally, that we write Brown's second relation between the parameters θ_0 and ρ_0 in the form

$$\theta_0 = F(\rho_0). \quad (29)$$

Put

$$f(r, \theta, \rho_0) \equiv r - \rho_0 \exp \left[\left(\frac{\rho_0^3}{\gamma_0 M t_0^2} \right)^{\frac{1}{2}} (\theta - F(\rho_0)) \right]. \quad (30)$$

Then by (21), the orbits considered are the family

$$f(r, \theta, \rho_0) = 0, \quad (31)$$

and their envelope is obtained by eliminating ρ_0 between (31) and

$$\frac{\partial f}{\partial \rho_0} = 0. \quad (32)$$

Now if the orbit-density *along the present positions* $\theta_0 = F(\rho_0)$ is given as $n_0(\rho_0)d\rho_0$, the corresponding *radial* orbit-density at the time when the radius vector was r was $n(r)dr$, where as before

$$n(r)dr = n_0(\rho_0)d\rho_0$$

and the differentials dr , $d\rho_0$ are connected by

$$\frac{\partial f}{\partial r} dr + \frac{\partial f}{\partial \rho_0} d\rho_0 = 0.$$

Thus

$$n(r) = -n_0(\rho_0) \frac{\partial f / \partial r}{\partial f / \partial \rho_0} = \frac{n(\rho_0)}{-\partial f / \partial \rho_0}, \quad (33)$$

which again, by (32), tends to infinity near the point of contact with the envelope.

15. Brown stated his hypothesis in the form that* "the arms are the envelopes of the paths followed by the particles or masses. From this point of view, the visible structure is due to the greater space-density of visible matter in the neighbourhood of the arms than elsewhere". Our analysis confirms this statement. But further consideration shows that the point of contact of each orbit with its envelope will not be occurring at the present moment.

In the general case (29), the polar co-ordinates of the actual point of contact are found to be given by

$$\theta = -\frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} + F(\rho_0) + \frac{2}{3} \rho_0 F'(\rho_0), \quad (34)$$

$$r = \rho_0 \exp \left[-\frac{2}{3} + \frac{2}{3} \left(\frac{\rho_0^3}{\gamma_0 M t_0^2} \right)^{\frac{1}{2}} \rho_0 F'(\rho_0) \right]. \quad (35)$$

This position will coincide with the present position

$$\theta = F(\rho_0), \quad r = \rho_0,$$

only if

$$\rho_0 F'(\rho_0) = \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}}.$$

Integrating, this requires

$$F(\rho_0) = -\frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} + \text{const.} \quad (36)$$

Suppose this relation holds good at time t_0 . This means that the present positions are given by

$$\theta_0 = F(\rho_0) = -\frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} + C. \quad (37)$$

* *Loc. cit.*, p. 110.

Inserting this value of θ_0 in the parametric equations for the orbit, (19) and (20), and using (20) to eliminate ρ_0 , we see that the positions of the particles at any other epoch t lie along the path

$$\theta + \frac{2}{3} \left(\frac{\gamma_0 M t^3}{r^3 t_0} \right)^{\frac{1}{2}} - C = \left(\frac{\gamma_0 M t^3}{r^3 t_0} \right)^{\frac{1}{2}} \log \frac{t}{t_0}$$

or

$$\theta = C + \left(\frac{\gamma_t M t^2}{r^3} \right)^{\frac{1}{2}} \left[\log \frac{t}{t_0} - \frac{2}{3} \right]. \quad (38)$$

On the other hand, since t_0 may equally well be considered an arbitrary epoch, the positions at time t should be derivable from (37) by changing θ_0 into θ , ρ_0 into r , t_0 into t and γ_0 into γ_t . Hence the t -positions should be given by

$$\theta = - \frac{2}{3} \left(\frac{\gamma_t M t^2}{r^3} \right) + C.$$

Formulae (38) and (39) are discrepant. Accordingly we cannot arrange for the present positions of the particles to be the points of contact of the orbits with their envelope.

This means that the present positions of the particles—that is, the spiral arms of a real nebula—cannot permanently be the points of contact of the orbits with their envelope. Moreover, the orbits cannot have had contact with an envelope in the past without having been seriously perturbed by their proximity to one another—that is to say, had they had an envelope in the past, they would be unlikely to exhibit now the regularity of present position we see in the spiral arms.

16. In fact, Brown's hypothesis is not only inadequate, but the exact opposite of what must be the actual situation. The orbits cannot have had an envelope! On the other hand, the existence of definite spiral arms in nebulae suggests that the present positions are the results of motion in a one-parameter family of orbits. We have therefore to consider what one-parameter family of orbits will escape having an envelope.

17. We continue to assume that the orbits lie in a plane and possess the same sense. Then relations (34) and (35) always determine an envelope unless the right-hand side of (34) happens to vanish identically. Let us examine this possibility. It would mean, *prima facie*, that the envelope reduced to, or was part of, the radius vector $\theta = 0$. Since the origin of θ is arbitrary, we can gain greater generality by putting the right-hand side of (34) equal to a constant, say β_0 , independent of ρ_0 . We get then

$$\frac{2}{3} \rho_0 F'(\rho_0) + F(\rho_0) - \frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} = \beta_0, \quad (40)$$

or

$$\frac{2}{3} \frac{d}{d\rho_0} [\rho_0^{\frac{3}{2}} F(\rho_0)] - \frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{\rho_0^2} \right)^{\frac{1}{2}} = \beta_0 \rho_0^{\frac{1}{2}}.$$

Integrating and dividing by $\rho_0^{\frac{3}{2}}$, we get

$$F(\rho_0) = \beta_0 + \frac{\text{const.}}{\rho_0^{\frac{3}{2}}} + \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \log \rho_0. \quad (41)$$

Without loss of generality we can write this in the form

$$F(\rho_0) \equiv \beta_0 + \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \log \frac{\rho_0}{\rho_1}, \quad (42)$$

where ρ_1 is arbitrary. The present positions of the particles, being $\theta_0 = F(\rho_0)$, are now given by

$$\theta_0 = \beta_0 + \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \log \frac{\rho_0}{\rho_1}. \quad (43)$$

If an envelope now exists, it can only lie along the radius vector

$$\theta = \beta_0. \quad (44)$$

Relation (35) then gives for the point of contact with this envelope, if it exists,

$$r = \rho_1. \quad (45)$$

But ρ_1 is a constant. Hence all the orbits now pass through a fixed point (ρ_1, β_0) , and the envelope reduces to a single point.

17. This suggests that we should now find directly the present positions (ρ_0, θ_0) of the particles which have passed through the point (ρ_1, β_0) at various times in the past. By (20) and (19), these present co-ordinates (ρ_0, θ_0) must satisfy, for some value of t , say $t = t_1$, the relations

$$\rho_1 = \rho_0 \frac{t_1}{t_0}, \quad \beta_0 - \theta_0 = \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \log \frac{t_1}{t_0}. \quad (46), (47)$$

Eliminating t_1 , the present positions (ρ_0, θ_0) must satisfy

$$\theta_0 = \beta_0 + \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \log \frac{\rho_0}{\rho_1}, \quad (48)$$

which is just (43).

Thus we have reached the conclusion that if the present existence of nebular arms implies that the orbits have had no envelope in the past, then the orbits must have originated from a point $r = \rho_1, \theta = \beta_0$. This strikingly reminds us of Jeans's suggestion that spiral arms are the consequences of tidal emissions from fixed points. It is also in conformity with Lindblad's conclusion that on his "instability" theory the points of ejection are fixed in space.

17 a. We note that ρ_1 is the radial co-ordinate of the point of ejection in t -measure. In τ -measure, the radial co-ordinate ρ of the point of ejection is given, by (20), by

$$\rho = \rho_1 \frac{t_0}{t_1}. \quad (48 a)$$

As t_1 increases, ρ steadily decreases. Thus in τ -measure, the point of ejection steadily moves inwards along the radius vector $\theta = \beta_0$. A referee has kindly pointed out to me that (48) can readily be derived by using this circumstance and employing τ -measure. The particle ejected at time τ_1 moves in τ -measure in a circle of radius ρ with constant angular velocity $(\gamma_0 M / \rho^3)^{\frac{1}{2}}$. The angular displacement in the interval τ_1 to t_0 , when added to β_0 , gives the present angular position as

$$\theta = \beta_0 + (\gamma_0 M / \rho^3)^{\frac{1}{2}} (t_0 - \tau_1). \quad (48 b)$$

But

$$\tau_1 = t_0 \log (t_1 / t_0) + t_0.$$

Eliminating τ_1 from (48 *b*) in terms of ρ_1 given by (48 *a*), we get for the present position

$$\theta = \beta_0 + \left(\frac{\gamma_0 M t_0^2}{\rho^3} \right)^{\frac{1}{2}} \log \frac{\rho}{\rho_1}$$

in agreement with (48). The referee's method indicates very clearly how the present spiral arms are reconcilable with the description of circular orbits in τ -measure.

18. *The functional equation satisfied by a nebular arm.* To be quite fair, we must now apply the criterion of Section 15, whose non-satisfaction was there one of the reasons compelling us to reject Brown's hypothesis. We want to test whether solution (43) or (48), for the *present* positions of the particles, represents a permanent or an ephemeral state of affairs.

We have hitherto described the present positions of the particles by the equation $\theta_0 = F(\rho_0)$. Let us now represent the positions of the particles at *any* time t by the equation

$$\theta = F(r, t), \quad (49)$$

so that the present positions are represented by

$$\theta_0 = F(\rho_0, t_0). \quad (50)$$

The separate orbits through points (ρ_0, θ_0) being given by (19) and (20), there must be a functional equation satisfied by F , supposed at this stage not yet known. This we proceed to find.

By displacement along the orbit (19), the present position (ρ_0, θ_0) becomes at time t the position

$$\theta = F(\rho_0, t_0) + \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \log \frac{t}{t_0}, \quad (51)$$

$$r = \rho_0 t / t_0. \quad (52)$$

Substituting for ρ_0 in terms of r , this position becomes

$$\theta = F\left(\frac{rt_0}{t}, t_0\right) + \left(\frac{\gamma_0 M t^3}{r^3 t_0}\right)^{\frac{1}{2}} \log \frac{t}{t_0},$$

or, using (12) and (13),

$$\theta = F\left(\frac{rt_0}{t}, t_0\right) + \left(\frac{\gamma_t M t^2}{r^3}\right)^{\frac{1}{2}} \log \frac{t}{t_0}. \quad (53)$$

This must be identical with $\theta = F(r, t)$. Accordingly

$$F\left(\frac{rt_0}{t}, t_0\right) + \left(\frac{\gamma_t M t^2}{r^3}\right)^{\frac{1}{2}} \log \frac{t}{t_0} \equiv F(r, t). \quad (54)$$

This is satisfied by

$$F(r, t) \equiv \beta_0 + \left(\frac{\gamma_t M t^2}{r^3}\right)^{\frac{1}{2}} \log \frac{r}{\rho_1}. \quad (55)$$

For the left-hand side of (54) is then

$$\beta_0 + \left[\gamma_0 M t_0^2 \left(\frac{t}{rt_0} \right)^3 \right]^{\frac{1}{2}} \log \frac{rt_0}{\rho_1 t} + \left(\frac{\gamma_t M t^2}{r^3} \right)^{\frac{1}{2}} \log \frac{t}{t_0},$$

which reduces to (55).

To obtain the general solution of the functional equation (54), put

$$F(r, t) = \beta_0 + \left(\frac{\gamma_t M t^2}{r^3} \right)^{\frac{1}{2}} \log \frac{r}{\rho_1} + \psi(r, t). \quad (56)$$

Then we find

$$\psi\left(\frac{rt_0}{t}, t_0\right) = \psi(r, t), \quad (57)$$

the general solution of which is

$$\psi(r, t) = \psi\left(\frac{t}{r}\right). \quad (58)$$

Thus the general solution of the nebular functional equation (54) is

$$F(r, t) = \beta_0 + \left(\frac{\gamma_t M t^2}{r^3} \right)^{\frac{1}{2}} \log \frac{r}{\rho_1} + \psi\left(\frac{t}{r}\right). \quad (59)$$

I shall not here pursue the consequences of the more general solution (59), but be content with the physically significant solution (55). This gives as the *permanent* equation of the nebular arm at time t the locus

$$\theta = \beta_0 + \left(\frac{\gamma_t M t^2}{r^3} \right)^{\frac{1}{2}} \log \frac{r}{\rho_1}, \quad (60)$$

where β_0 and ρ_1 are arbitrary and γ_t denotes the value of the "constant" of gravitation at time t , by (12). Putting $t = t_0$, $r = \rho_0$, $\theta = \theta_0$ and $\gamma_t = \gamma_0$, this reduces to (48).

19. We have now obtained an equation for the locus, at any time t , of the positions of the particles comprising a nebular arm by two distinct methods: one being the expression of the non-existence of an envelope, the other through a functional equation. They gave the same result; and we have interpreted this result physically as being the present positions of particles emitted in the past from a fixed point in the nebula. Let us consider the nature of the locus at time t_0 , as given by (43) or (48), in which ρ_0 , θ_0 are now to be considered as current polar co-ordinates. It must be borne in mind that the fact that (60) reduces to (48) on setting $t = t_0$ depends in an essential way on the secular dependence of γ on t .

20. First, let us determine the time t_1 of emission of the particle now at (ρ_0, θ_0) from the point of ejection (ρ_1, β_0) . This is given, by (20) or (52), by

$$t_1 = t_0 \frac{\rho_1}{\rho_0}. \quad (61)$$

As ρ_0 increases, the associated value of t_1 decreases; in other words, the particles most distant from the nucleus were emitted the longest time ago.

21. The locus (48) is a spiral, in which ρ_0 varies with θ_0 . It is quite distinct from the spirals constituting the orbits. The latter are given by (21), in which (r, θ) are current variables, (ρ_0, θ_0) constants. In (48), on the other hand, though it is superficially of the same form, (ρ_0, θ_0) are current variables, (ρ_1, β_0) constants. The relation between the two is akin to the relation between pole and polar for a conic or other plane curve.

The locus (48) passes through the point (ρ_1, β_0) and again through the point (∞, β_0) . It therefore has $\theta = \beta_0$ as an asymptote. As ρ_0 increases from ρ_1 , θ_0 at first increases from β_0 , attains a maximum and then decreases again to β_0 .

There are thus a *finite* number of convolutions of the spiral in one sense, and then a *finite* number back again in the opposite sense, the first sense being that of the orbital motion. To see this, we have that the angle ϕ between tangent and radius vector is given by differentiation of (48) in the form

$$\tan \phi = \rho_0 \frac{d\theta_0}{d\rho_0} = \left(\frac{\gamma_0 M t_0^2}{\rho_0^3} \right)^{\frac{1}{2}} \left[1 - \frac{3}{2} \log \frac{\rho_0}{\rho_1} \right]. \quad (62)$$

This vanishes, and θ_0 has a maximum, when

$$\rho_0 = \rho_1 e^{\frac{2}{3}} = 1.9477 \rho_1, \quad (63)$$

in which case

$$\theta_0 = \beta_0 + \frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{e^2 \rho_1^3} \right)^{\frac{1}{2}}. \quad (64)$$

The particle *now* at this point of maximum θ_0 was emitted, by (61), at epoch

$$t_1 = t_0 e^{-\frac{2}{3}} = 0.5134 t_0. \quad (65)$$

All particles beyond this point were emitted at earlier times t_1 , all particles interior to this point at later times t_1 .

The spiral locus executes a finite number of convolutions, in the sense of the orbital motion, between the point of emission $\rho_0 = \rho_1$, $\theta_0 = \beta_0$ and the point of contact of the locus with $\theta_0 = (\theta_0)_{\max}$, and thereafter an equal number in the opposite sense, becoming ultimately asymptotic to $\theta_0 = \beta_0$. This finite number of convolutions, ν_0 is given, by (64), by

$$\nu_0 = \frac{1}{2\pi} \cdot \frac{2}{3} \left(\frac{\gamma_0 M t_0^2}{e^2 \rho_1^3} \right)^{\frac{1}{2}}. \quad (66)$$

22. For a galaxy comparable with our own galaxy, since M/ρ_1^3 is a measure of the density of matter inside the distance ρ_1 , we may suppose M/ρ_1^3 to have an order of magnitude given by J. S. Plaskett's values*

$$M = 1.65 \times 10^{11} \odot, \quad \rho_1 = 10^4 \text{ parsecs} = 3.08 \times 10^{18} \times 10^4 \text{ cm.}$$

Taking $t_0 = 2 \times 10^9$ years $= 0.6 \times 10^{17}$ secs., the order of magnitude of ν_0 is then given by

$$\nu_0 = \frac{1}{3\pi} \left[\frac{6.66 \times 10^{-8} \times 1.65 \times 10^{11} \times 2 \times 10^{33} \times 0.36 \times 10^{34}}{2.718^2 \times 3.08^3 \times 10^{54} \times 10^{12}} \right]^{\frac{1}{2}} = 2.0.$$

It is well known that spiral nebulae exhibit a finite number (usually about two †) of convolutions, but I am very anxious not to press unduly the agreement we have found. It is for this reason that I have confined myself to J. S. Plaskett's fairly well-determined values of M and ρ_1 , rather than attempt to estimate them for individual external galaxies. But the calculation is very striking, considering the diversity of the data brought together.

For a nebula like the Andromeda nebula, M 31, both the nuclear mass M and the probable value of ρ_1 may be very much smaller than we have taken for our galaxy.‡ Indeed, if we accept Hubble's estimate of the nuclear mass as $2.4 \times 10^8 \odot$ inside 200 parsecs, the value of M/ρ_1^3 is altered from $(1.65 \times 10^{11})/(10^4)^3$ to

* Halley Lecture, "The Dimensions and Structure of the Galaxy", p. 22, Oxford, 1935.

† E.g., Jeans, *Astronomy and Cosmogony*, p. 351.

‡ Cf. Hubble, *The Realm of the Nebulae*, p. 134; or *Contr. Mt. Wilson Obs.*, No. 376; *Ap. J.*,

$(2.4 \times 10^8)/(2 \times 10^2)^3$, *i.e.* from 0.165 to 30, and the computed value of ν_0 is increased in the ratio $(30/0.165)^{\frac{1}{3}}$ or 13.5 times. It is plausible to suppose that the actual number of convolutions in the unresolved region exterior to the nucleus proper may be quite large.

23. But before we can subject our analysis to detailed numerical tests, it is necessary to be clear about its physical interpretation. The analysis has given a model in which the orbits are equi-angular spirals, and in which the *present positions* of the particles lie in spiral arms (not equi-angular spirals) which unwind at first in the sense of the orbital motion for a finite number ν_0 of convolutions between $\rho_0 = \rho_1$ and $\rho_0 = e^{\frac{2}{3}}\rho_1 = 1.95\rho_1$ and thereafter unwind backwards between $\rho_0 = 1.95\rho_1$ and $\rho_0 = \infty$, ρ_1 being the radius vector of the point of emission. As the diagrams will show, the turn-back is sharper the larger is ν_0 , and indeed

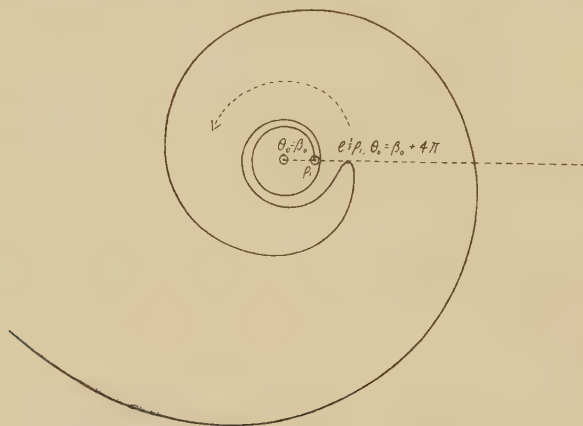


FIG. 2.—Graph of $\theta_0 = \beta_0 + 3\pi\nu\left(\frac{\rho_1}{\rho_0}\right)^{\frac{2}{3}} \log \frac{\rho_0}{\rho_1}$, with $\nu = 2$, (ρ_0, θ_0) being polar co-ordinates.

The locus of (ρ_0, θ_0) gives the predicted present positions of the particles forming the nebula. These particles have been emitted from (ρ_1, β_0) . The arrow gives the sense of the orbital motion.

quite sharp for ν_0 as small as 2. I suggest that this doubling back of the spiral arms in our model may be the origin of the discrepancy between Hubble's and Lindblad's conclusions as to the sense of unwinding of the spiral arms.* In the outer parts of the nebula the sense of the arms should be retrograde as compared with the orbital motion; in the parts closer to the nucleus the sense of the arms should be direct. It should be remembered that there is no necessity for an actual spiral to show the complete set of convolutions, both direct and retrograde; ejection which occurred earlier than $t_1 = 0.5134t_0$ will be represented only by the outer retrograde convolutions, and so if ejection ceased at some epoch earlier than $t_1 = 0.5134t_0$, no inner convolutions will be represented. Likewise a more recent series of ejections, later than $t_1 = 0.5134t_0$, will be represented only by inner convolutions.

A further suggestion is that a "barred" spiral may be one for which a substantial part of the ejection took place in the vicinity of the epoch $t_1 = 0.5134t_0$, so that most of the arm is tangent to $\theta_0 = (\theta_0)_{\max}$. Fig. 3 below, which shows two arms arising from ejections at antipodal points for the case $\nu_0 = \frac{1}{2}$, strongly suggests certain forms of "barred" spirals.

* See, e.g., Hubble, *Ap. J.*, 97, 112, 1943.

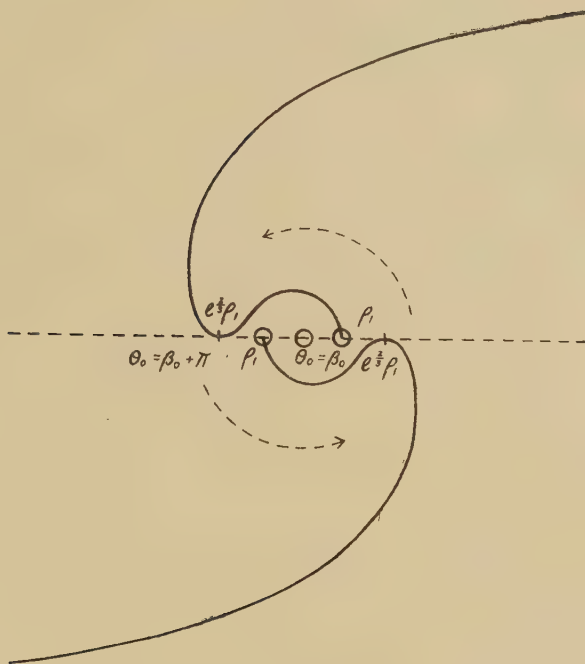


FIG. 3.—Graph of $\theta_0 = \beta_0 + 3\pi e\nu \left(\frac{\rho_1}{\rho_0}\right)^{\frac{2}{3}} \log \frac{\rho_0}{\rho_1}$ with $\nu = \frac{1}{2}$, (ρ_0, θ_0) being polar co-ordinates.

The locus (ρ_0, θ_0) gives the predicted present positions of the particles forming the nebula. These particles have been emitted from (ρ_1, β_0) . The arrows give the sense of the orbital motion.

24. Whatever may be the ultimate fate of these suggestions, it is clear that the quantity we have called ν_0 , which measures the number of convolutions in each of the two parts of the spiral, is the fundamental parameter of the theory. Given the numerical magnitude of ν_0 , we can determine M in terms of ρ_1 or ρ_1 in terms of M , at any epoch t_0 —that is to say, we can determine the size for given mass, once ν_0 is known.

Again, the equation for a nebular arm, (48), can be conveniently expressed in terms of ν_0 and the scale and orientation parameters, ρ_1 and β_0 . For, using (66), equation (48) may be written

$$\theta_0 = \beta_0 + 3\pi e\nu_0 \left(\frac{\rho_1}{\rho_0}\right)^{\frac{2}{3}} \log \frac{\rho_0}{\rho_1}. \quad (67)$$

This is the most convenient form for graphical and numerical calculations of the theoretical shape of a spiral arm.

Again, if we write

$$\nu = \frac{1}{2\pi} \cdot \frac{2}{3} \left(\frac{\gamma_t M t^2}{e^2 \rho_1^3} \right)^{\frac{1}{2}}, \quad (68)$$

then when t_0 changes to a value t , and (48) changes by the march of time into (60), (67) passes into

$$\theta = \beta_0 + 3\pi e\nu \left(\frac{\rho_1}{r}\right)^{\frac{2}{3}} \log \frac{r}{\rho_1}. \quad (69)$$

Remembering that ν is a function of t , we have here the most concise form of the permanent equation of a nebular arm. We notice from (66) and (68) that

$$\nu = \nu_0 \left(\frac{t}{t_0} \right)^{\frac{2}{3}}, \quad (70)$$

and so the number of convolutions in either "half" of a completed nebular arm varies as $t^{\frac{2}{3}}$.

Physically, the parameter ν , which depends on t and M/ρ_1^3 , describes the physical state of the nucleus from which ejection has occurred; it controls the *shape* of the nebular arm, just as ρ_1 controls its scale and β_0 its orientation.

25. Again, the "angle" of the spiral arm at any point is given by

$$\tan \phi = \rho_0 \frac{d\theta_0}{d\rho_0} = 3\pi e \nu_0 \left[1 - \frac{3}{2} \log \frac{\rho_0}{\rho_1} \right] \left(\frac{\rho_1}{\rho_0} \right)^{\frac{2}{3}}. \quad (71)$$

This shows immediately that $(\theta_0)_{\max}$, which is reached when $\log(\rho_0/\rho_1) = \frac{2}{3}$, takes the value

$$(\theta_0)_{\max} = \beta_0 + 3\pi e \nu_0 (e^{-1})^{\frac{2}{3}} = \beta_0 + 2\pi \nu_0. \quad (72)$$

At the point of ejection, namely $\rho_0 = \rho_1$, the value of the "angle" of the spiral is given by

$$\tan \phi = 3\pi e \nu_0 = \left(\frac{\gamma_0 M t_0^2}{\rho_1^3} \right)^{\frac{1}{3}} = [\tan \alpha]_{\theta_0 = \theta_1}, \quad (73)$$

where $\tan \alpha$ is the parameter of a spiral *orbit*, as given by (22). As an example, take $\nu_0 = 2$. Then $\tan \phi$ at the point of ejection is equal to $6\pi e$ or 51.2 , and $\phi = 88^\circ.9$. It is clear that near the point of ejection the convolution begins by being nearly circular.

The mean value of the tangent of the "angle" of a spiral arm in the inner region between $\rho_0 = \rho_1$ and $\rho_0 = e^{\frac{2}{3}}\rho_1$ is given by

$$\begin{aligned} \overline{\tan \phi} &= \frac{1}{2\pi \nu_0} \int_{\theta_0 = \beta_0}^{\theta_0 = \beta_0 + 2\pi \nu_0} \tan \phi d\theta_0 = \frac{3}{2} e \int_1^{e^{\frac{2}{3}}} \left[1 - \frac{3}{2} \log x \right] x^{-\frac{2}{3}} d(x^{-\frac{1}{3}} \log x) \\ &= \frac{3}{8} e - \frac{1}{8} e^{-1} = 0.9734, \end{aligned} \quad (74)$$

corresponding to a "mean" spiral "angle" of $44^\circ.2$ in this region. This is in fair accord with observation.*

26. Figs. 2 and 3 give graphs of theoretical spiral arms calculated from (67) for the cases $\nu_0 = 2$ and $\nu_0 = \frac{1}{2}$. In the latter case two antipodal arms are shown. It is of course not necessary to fix the value of ρ_1 absolutely; any convenient scale may be taken in which the unit represents ρ_1 . The sharpness of the turn-back in the case $\nu_0 = 2$ will be noticed, and also the ϕ -shape of the pair of arms for the case of ν_0 small ($\nu_0 = \frac{1}{2}$). Strictly speaking, these diagrams should be completed by the addition of the arcs of spiral *orbits*, each arc preceding from the point of ejection, (ρ_1, β_0) and terminating (at epoch t_0) on some particular point of the spiral arm. These orbit-spirals will be of varying angles, the angle diminishing, according to (22), as ρ_0 the present radial distance from the nucleus increases. For ρ_0 near ρ_1 , the orbits are represented by short arcs approximating

* Cf. Von der Pahlen, *Ast. Nach.*, 188, Nr. 4503, 1911; Groot, *M.N.*, 85, 535, 1925; Reynolds, *M.N.*, 85, 1014, 1925.

closely to the spiral arm itself; for ρ_0 very large, the orbit is represented almost by a rectilinear radius-vector, of very small spiral curvature.

27. The variation of line-density to be expected along spiral arms of these types is readily calculated. Let t_1 be the epoch of emission of the material now at ρ_0 , and let $n(t_1)dt_1$ give the rate of emission, of course at the fixed distance ρ_1 from the centre. Then as usual $\rho_1 = \rho_0(t_1/t_0)$. The particle emitted from ρ_1 at time t_1 arrives at ρ_0 at time t_0 ; the particle emitted from ρ_1 at the earlier time $t_1 - dt_1$ arrives at distance $\rho_0 + d\rho_0$ at time t_0 . Hence, since ρ_1 is fixed,

$$t_1 d\rho_0 - \rho_0 dt_1 = 0.$$

Now suppose that $N(\rho_0)d\rho_0$ is the present distribution of the particles in radius vector. Then

$$N(\rho_0)d\rho_0 = n(t_1)dt_1,$$

and so

$$N(\rho_0) = n(t_1) \frac{t_1}{\rho_0}.$$

But $t_1 = t_0(\rho_1/\rho_0)$. Hence

$$N(\rho_0) = \frac{n(t_1)t_0\rho_1}{\rho_0^2}. \quad (75)$$

The density of particles in radius vector, along an arm, should therefore fall off as the inverse square of the radial distance, as long as the rate of emission remains constant.

It is well known from the researches of Hubble and others that the distribution of luminosity in certain elliptical nebulae obeys an inverse-square law of the type (distance + const.)⁻², but I am not aware of any measures of density-variation along a spiral arm. All that can be said is that the law (75) is not in disaccord with general impressions.

Wadham College, Oxford:

1946 January.

Note added 1946 January 26.—Since the above was prepared for press, my attention has been directed to *Stockh. Obs. Ann.* **14**, No. 3, 1942—a reference kindly supplied by Dr A. Hunter—wherein Lindblad brings forward actual evidence that in the two spirals NGC 2681 and 3190 there occur faint outer arms which *trail*, in addition to the main inner arms which proceed in the sense of the rotation. This is in accordance with the theory developed in the present paper. Lindblad in the paper cited also suggested that the existence of a turning-point in the spiral arms, and consequently the co-existence in the same nebula of inner spiral arms proceeding in the direction of the rotation and fainter outer arms proceeding in the retrograde direction, could account for the contradictory results that have been obtained by different investigators on the sense of unwinding of the spiral arms in relation to the sense of the rotation.* I was quite unaware of this suggestion of Lindblad's when my kinematic investigations led me to the same conclusion. Lindblad has also suggested that a similar state of affairs may exist in the Andromeda nebula, M31. The actual gravitational analysis by which Lindblad concludes the existence of a doubling back of a spiral arm is different from that in the present paper.

* See Hubble, *Ap. J.*, **97**, 112, 1943.

MEAN PARALLAX DETERMINATIONS FROM PROPER MOTIONS

Archibald Brown

(Received 1946 April 1).

Summary

The paper records mean parallaxes of stars of apparent magnitude 7.0 to 10.0 (photographic) sub-divided according to spectral type and galactic latitude. The parallaxes were derived from proper motions of stars in declination -40° to -52° , using the total proper motions of the stars and postulating that the frequency distribution of the velocities of the stars concerned can be represented by the two-stream theory.

1. *Introduction*.—Van Rhijn and Bok* have combined results from a number of catalogues of proper motions to give determinations of mean parallaxes for stars of different apparent magnitude and galactic latitude. Their work has become a standard of reference to such an extent that it will be convenient to describe the present paper, although much more restricted in its material, in terms of their results and methods. The present paper gives determinations of mean parallaxes from a single catalogue†, over a limited range of magnitudes (namely 7.0 to 10.0 apparent magnitude). It differs from the Groningen investigation mainly in separating the material into three groups according to spectral type and in using the total proper motion of each star instead of the component of proper motion in the direction of the solar antapex.

The values of mean parallax obtained show the expected increase with increasing galactic latitude more markedly than the Groningen values and also indicate that comparatively few dwarf stars occur among the K- and M-type stars in the catalogue. Sections 2 to 5 of the paper give the methods of analysing the material; Sections 6 to 9 explain the tables and graphs in which the results are presented and compare them with results from other investigations.

2. *Observational Material*.—In a sense the results are not completely independent of the Groningen values since the observed proper motions relative to fainter reference stars were reduced to absolute proper motions by corrections for the parallactic motion of the reference stars and for the effect of galactic rotation taken from the Groningen tables. However, the reference stars were of 11.9 mean apparent magnitude and a comparison with Boss stars‡ gave no indication of systematic errors from this source. Details of the plates used and the reduction of the observations are given by Dr Jackson in an introduction to the Cape Catalogue. All stars down to magnitude 9.0 on the Cape Photographic Durchmusterung Scale were measured, but in some areas of the zone the measurements are complete for fainter stars. (This point re-appears later in discussing the

* *Publ. astr. Lab. Groningen*, No. 45, 1931.

† Cape Catalogue 2. Note: references in this paper to the Star Catalogues of the Royal Observatory, Cape of Good Hope, are given as follows:—

Cape Catalogue 1 = *Magnitudes of stars contained in the Cape Zone Catalogue of 20,843 stars, 1927.*

Cape Catalogue 2 = *Proper Motions of Stars in the Zone Catalogue of 20,843 stars, 1936.*

Cape Catalogue 3 = *Catalogue of 20,554 faint stars in the Cape Astrographic Zone, 1939.*

‡ Cape Catalogue 2, xxi.

results.) The zone of declination within which the stars lie passes about 10° from the southern galactic pole at one point (near R.A. 0^h) and elsewhere crosses the galactic equator, making it suitable for determining the variation of mean parallax with galactic latitude, although the variation in galactic longitude is small. For the investigation the zone was divided into 16 regions of right ascension 0^h to $1\frac{1}{2}^h$, $1\frac{1}{2}^h$ to 3^h , ... $22\frac{1}{2}^h$ to 24^h , using the full range in declination in each region. This follows Smart and Tannahill*, and in order to make use of their analysis of the motions according to spectral type the same grouping of spectral types was also used, viz. A5 to F5, F8 to G5 and K0 to M, neglecting the early type stars O, B and A0 to A3.† It will be seen from Dr Jackson's table of the distribution of stars with regard to spectral type (on p. xxii of his Introduction) that M-type stars form less than 10 per cent. of the K0 to M group. Magnitude groups were formed from 6.9 to 7.9, 8.0 to 8.9, 9.0 to 9.4, 9.5 to 9.8, 9.9 to 10.3, and in a few cases 10.4 to 11.2 apparent photographic magnitude. It was recognized that these groups would not all be complete, but provided there was no systematic selection of stars with large proper motions or stars with little apparent proper motion the mean parallaxes would still be of value. They were therefore included despite the knowledge that they might have to be discarded later.

3. *Analysis of Observations.*—The method adopted in evaluating the mean parallaxes was one given by Smart.‡ From the catalogue values of

$\mu_\alpha \cos \delta$ = proper motion in R.A. (in seconds of arc per annum),

μ_δ = proper motion in declination (in seconds of arc per annum),

the total proper motion

$$\mu = \{(\mu_\alpha \cos \delta)^2 + \mu_\delta^2\}^{\frac{1}{2}}$$

was tabulated for each star.§ Now, on the two-stream theory of stellar motions the mean transverse motion for a small region of the sky is a known function of the drift constants, and these constants can be determined from the directions of the motions in the region. In this case the motions were analysed by Smart and Tannahill|| and the required constants are given in Table I of their paper. Knowing the mean angular motion $\bar{\mu}$ and the mean transverse motion as above, we get the mean parallax from their ratio.¶ The procedure was therefore to evaluate $\bar{\mu}$ for each group of stars and to calculate the divisor D representing the mean transverse motion for each region from Smart and Tannahill's analysis. Table I shows the co-ordinates of the centres of the regions, the number of stars used in the analysis, and the values of D for the three spectral divisions.

Assuming the solar motion to be 19.5 km./sec., the equation for the mean parallax $\bar{\pi}$ becomes

$$\bar{\pi} = 0.242 \bar{\mu} / D, \quad (1)$$

and this is the fundamental formula used in the investigation.

* *M.N.*, 100, 688, 1940.

† Stars with composite spectra were also ignored. For double stars, the spectral type and magnitude of the brighter component were used where details were given separately.

‡ *Stellar Dynamics*, pp. 208–217, Cambridge, 1938.

§ For double stars the proper motion of the centre of mass was used where given and the mean of the proper motions in other cases.

|| *Loc. cit.*

¶ There is an ellipsis in the argument here, but the outline is clear. Strictly, the ratio depends on the distribution of parallax and transverse velocity.

TABLE I
Position of Centre of Region and Number of Stars
l=galactic longitude; *b*=galactic latitude

Region	Co-ords. of centre			Number of Stars			Value of factor <i>D</i>			Mag- nitude correc- tion <i>δm</i>
	R.A.	<i>l</i>	<i>b</i>	A5-F5	F8-G5	Ko-M	A5-F5	F8-G5	Ko-M	
	h m	°	°							
1	0 45	267.1	-72.2	124	266	176	1.425	1.355	1.194	-0.1
2	2 15	229.9	-64.2	119	234	145	1.410	1.382	1.369	-0.2
3	3 45	218.9	-49.7	145	271	187	1.445	1.342	1.261	-0.2
4	5 15	218.3	-34.0	190	247	215	1.275	1.348	1.245	-0.4
5	6 45	222.6	-18.8	265	197	249	1.553	1.347	1.263	+0.05
6	8 15	230.2	- 5.1	219	188	276	1.549	1.424	1.439	-0.2
7	9 45	241.0	+ 6.2	250	190	290	1.985	1.529	1.621	-0.15
8	11 15	254.9	+13.7	196	197	317	1.748	1.604	1.604	0.0
9	12 45	270.9	+16.2	211	280	341	1.915	1.800	1.646	0.05
10	14 15	286.8	+13.1	228	281	297	1.790	1.556	1.597	0.05
11	15 45	300.4	+ 5.1	215	258	307	1.971	1.735	1.469	-0.1
12	17 15	310.8	- 6.5	202	213	268	1.645	1.706	1.717	+0.2
13	18 45	318.1	-20.5	228	266	252	1.531	1.435	1.424	-0.25
14	20 15	321.9	-35.8	235	258	222	1.421	1.318	1.507	0.0
15	21 45	320.6	-51.4	163	273	227	1.307	1.380	1.212	-0.25
16	23 15	307.6	-65.5	148	276	236	1.344	1.295	1.384	-0.05
Total				3138	3895	4005				

4. *Effect of Errors in Proper Motions.*—Unfortunately a correction to $\bar{\mu}$ is required before equation (1) can be used, because of accidental errors in observing $\mu_{\alpha} \cos \delta$ and μ_{δ} . If we assume each component is determined subject to a probable error *r*, then the effect of the error is to give a biased estimate of $\bar{\mu}$. If we considered only $\mu_{\alpha} \cos \delta$, the average value $\overline{\mu_{\alpha} \cos \delta}$ would not be biased nor, similarly, would the average value of μ_{δ} be biased, but with the total proper motion μ , equal positive and negative errors in the components $\mu_{\alpha} \cos \delta$ and μ_{δ} do not cancel out. This can be seen from Fig. 1, in which the point P(*x*, *y*) represents the true value of

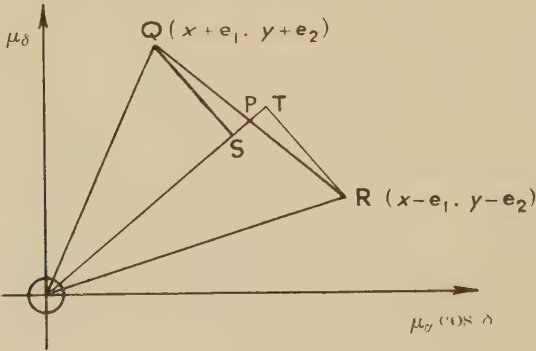


FIG. 1.

a proper motion and Q and R are the points $(x + e_1, y + e_2)$ and $(x - e_1, y - e_2)$ respectively, while QS and RT are perpendiculars from Q and R on OP. Then

$$2OP = OS + OT < OQ + OR,$$

so assuming a frequency distribution of the errors e which is symmetrical about zero the observed $\bar{\mu}$ will be greater than the true value. If PQ is small compared with OP in Fig. 1 the effect of this bias will be small, but if PQ is of the same magnitude or greater than OP the effect becomes considerable. Thus the magnitude of the correction to the observed value of $\bar{\mu}$ depends on the frequency distribution of the values of μ and on the magnitude of the errors, *i. e.* on the probable error r .

A working hypothesis for evaluating the correction c to be applied to the observed value of $\bar{\mu}$ is given by Smart*, with appropriate tables, and his method was used. The magnitude of the correction can be seen from Table II; it is as much as 10 per cent. of the observed value in some cases. Thus denoting

$\bar{\mu}$ = true value of mean total proper motion,

$\bar{\mu}_0$ = observed value of mean total proper motion,

$$\bar{\mu} = \bar{\mu}_0 + c = \bar{\mu}_0 - c', \quad (2)$$

where $c' = -c$ is shown in Table II. The value of $\bar{\mu}$ obtained in this way was used in equation (1) to obtain $\bar{\pi}$ (Table II).

TABLE II

Mean Proper Motions in each Group

Notes.—Magnitude Groups : a 6.9–7.9 d 9.5–9.8
 b 8.0–8.9 e 9.9–10.3
 c 9.0–9.4 f 10.4–11.2

$\bar{m}$ = Mean Apparent (Photographic) Magnitude.

N = No. of stars used. (The figure in brackets denotes the number omitted from the analysis. Where there is not a bracket, all the stars were used.)

$\bar{\mu}$ = Mean Proper Motion (corrected).

c' = Correction to observed mean proper motion.

Region	Spectral Type A5–F5				Spectral Type F8–G5				Spectral Type Ko–M			
	$\bar{m}$	N	$\bar{\mu}$	c'	$\bar{m}$	N	μ	c'	$\bar{m}$	N	$\bar{\mu}$	c'
1 a	7.52	12	53.59	1.08	7.68	9	179.6	0.54	7.51	12	57.45	0.80
b	8.51	33	52.07	1.23	8.52	35 (1)	94.02	0.64	8.38	26 (1)	36.58	1.57
c	9.20	28	41.39	1.32	9.22	33 (3)	60.86	1.20	9.21	29 (1)	24.38	1.96
d	9.65	22	32.19	1.76	9.67	56 (1)	62.61	1.14	9.63	32 (1)	24.45	1.96
e	10.08	29	32.78	1.50	10.07	85 (1)	47.06	1.29	10.09	40 (1)	26.02	1.96
f	—	—	—	—	10.59	42	46.94	1.39	10.61	32 (1)	24.45	1.96
2 a	7.61	10	48.90	0.80	7.64	17 (1)	102.65	0.64	7.32	8	42.67	1.08
b	8.59	44	46.51	1.08	8.50	34 (1)	66.89	1.20	8.54	37 (1)	26.78	1.79
c	9.18	26	32.75	1.37	9.23	34 (1)	55.69	1.25	9.25	26	28.83	1.67
d	9.66	20	44.69	1.26	9.65	52	56.20	1.32	9.69	14	18.95	2.26
e	10.11	19	29.44	1.67	10.09	60 (2)	42.51	1.44	10.10	35 (4)	25.12	1.79
f	—	—	—	—	10.67	32	47.30	1.39	10.71	20	23.30	3.05

* *Loc. cit.*, pp. 210–215.

TABLE II (cont.)

Region	Spectral Type A ₅ -F ₅				Spectral Type F ₈ -G ₅				Spectral Type K ₀ -M			
	$\bar{m}$	N	$\bar{\mu}$	c'	$\bar{m}$	N	$\bar{\mu}$	c'	$\bar{m}$	N	$\bar{\mu}$	c'
3 <i>a</i>	7.51	9	48.14	0.97	7.53	18	89.70	0.80	7.51	12	46.84	1.08
<i>b</i>	8.46	56	42.56	1.32	8.56	41	69.44	0.97	8.58	36	20.37	1.96
<i>c</i>	9.22	29	30.80	1.92	9.23	46	57.56	1.20	9.17	34 (1)	24.83	1.96
<i>d</i>	9.62	27	26.15	1.92	9.69	75 (2)	51.20	1.29	9.65	31 (2)	21.09	2.56
<i>e</i>	10.02	24	26.20	3.05	10.06	70 (1)	44.62	1.39	10.09	48	21.17	2.56
<i>f</i>	—	—	—	—	10.56	18	35.32	2.74	10.55	23	19.64	2.58
4 <i>a</i>	7.59	18	52.81	1.08	7.57	7	115.1	0.74	7.50	12 (1)	27.88	2.45
<i>b</i>	8.54	51	41.26	1.37	8.56	40 (2)	67.92	0.86	8.53	38	24.51	1.96
<i>c</i>	9.19	48	29.66	1.67	9.23	46 (1)	58.96	1.08	9.22	34	26.34	1.92
<i>d</i>	9.64	41	26.41	1.79	9.66	73 (1)	38.65	1.32	9.66	59 (2)	20.97	2.45
<i>e</i>	10.03	32	24.14	2.17	10.07	69	36.57	1.76	10.08	49	19.60	3.05
<i>f</i>	—	—	—	—	10.55	8	33.66	1.96	10.57	19 (1)	18.11	3.05
5 <i>a</i>	7.49	18	67.03	1.08	7.54	14 (1)	127.8	0.74	7.49	13 (1)	54.55	1.37
<i>b</i>	8.57	83	27.05	2.64	8.55	42 (1)	53.88	1.57	8.49	67 (2)	27.21	1.92
<i>c</i>	9.22	74	29.09	1.92	9.23	47	59.31	1.22	9.21	59 (4)	16.93	2.10
<i>d</i>	9.66	60	25.46	1.96	9.64	60 (2)	30.73	2.17	9.66	78 (2)	19.66	2.26
<i>e</i>	10.01	30	17.99	2.58	10.01	30	49.80	1.57	10.01	23	17.04	2.92
<i>f</i>	—	—	—	—	—	—	—	—	—	—	—	—
6 <i>a</i>	7.49	23	51.64	0.97	7.52	15	87.13	0.74	7.57	36	29.66	1.92
<i>b</i>	8.56	71	32.15	1.68	8.55	58	52.65	1.57	8.51	89 (1)	20.83	1.96
<i>c</i>	9.19	61 (1)	19.80	2.56	9.18	65	45.30	1.33	9.20	75 (1)	19.91	2.26
<i>d</i>	9.65	43	19.14	2.26	9.62	38	38.08	1.55	9.64	60	16.57	3.05
<i>e</i>	10.00	20	22.30	2.50	9.96	12	27.29	1.96	9.99	13 (1)	27.25	1.67
<i>f</i>	—	—	—	—	—	—	—	—	—	—	—	—
7 <i>a</i>	7.64	29	43.33	1.57	7.66	16	68.06	1.38	7.50	31	50.08	1.08
<i>b</i>	8.56	72 (1)	34.66	1.67	8.52	48	57.09	1.22	8.57	65 (2)	26.48	1.92
<i>c</i>	9.22	70	28.58	1.68	9.24	66	45.58	1.57	9.23	69 (1)	22.92	2.50
<i>d</i>	9.64	56	26.11	1.91	9.67	35	44.99	1.32	9.63	58	20.45	2.26
<i>e</i>	10.03	22	25.60	2.26	10.06	25	35.26	1.50	10.07	64	23.09	2.50
<i>f</i>	—	—	—	—	—	—	—	—	—	—	—	—
8 <i>a</i>	7.53	30	82.93	0.80	7.63	11	69.78	0.58	7.55	21	35.67	1.57
<i>b</i>	8.48	52	38.28	1.20	8.55	52	73.90	1.02	8.49	78	29.81	1.82
<i>c</i>	9.24	56	34.59	1.50	9.22	46	55.40	1.08	9.18	89 (1)	27.60	2.04
<i>d</i>	9.63	38	26.68	2.03	9.65	43 (1)	44.92	1.38	9.67	60 (1)	24.99	1.96
<i>e</i>	10.04	20	24.18	1.67	10.00	33 (1)	52.38	1.14	10.05	66 (1)	18.39	2.26
<i>f</i>	—	—	—	—	10.52	10	49.50	0.80	—	—	—	—
9 <i>a</i>	7.57	18	67.92	0.80	7.52	13	171.77	0.54	7.50	29 (1)	50.15	1.26
<i>b</i>	8.56	63	37.65	1.37	8.59	30	76.42	0.91	8.60	54	30.05	1.67
<i>c</i>	9.22	50	31.13	1.67	9.24	48 (1)	54.11	1.22	9.23	56 (3)	22.88	1.96
<i>d</i>	9.67	47	21.95	1.96	9.66	90 (1)	51.92	1.20	9.65	69 (3)	25.08	1.96
<i>e</i>	10.08	33	24.44	2.26	10.07	79 (2)	46.32	1.57	10.08	94	17.93	2.56
<i>f</i>	—	—	—	—	10.50	16	41.32	1.37	10.55	31 (1)	15.21	2.92

TABLE II (cont.)

Region	Spectral Type A5-F5				Spectral Type F8-G5				Spectral Type Ko-M			
	$\bar{m}$	N	$\bar{\mu}$	c'	$\bar{m}$	N	$\bar{\mu}$	c'	$\bar{m}$	N	$\bar{\mu}$	c'
10 a	7.48	15	92.16	0.64	7.48	14	108.55	0.74	7.58	24 (1)	47.92	1.08
b	8.62	54	35.78	1.37	8.57	48	76.16	0.99	8.58	64 (2)	34.83	1.55
c	9.18	54	30.33	1.50	9.22	54 (2)	52.52	1.02	9.20	69	22.04	2.26
d	9.63	56	27.35	1.67	9.68	71 (2)	38.94	1.32	9.64	52 (1)	27.54	1.85
e	10.05	49	23.37	1.96	10.06	79 (1)	43.64	1.32	10.06	81 (3)	21.61	2.45
f	—	—	—	—	10.49	10	27.93	2.07	—	—	—	—
11 a	7.52	21	51.78	1.08	7.71	18	71.20	0.80	7.46	31	43.58	1.26
b	8.51	47	36.57	1.37	8.53	46 (2)	70.57	0.91	8.53	54 (2)	27.76	1.96
c	9.19	63	27.58	1.67	9.25	57 (1)	51.96	1.22	9.23	70 (1)	22.71	2.45
d	9.65	50	22.84	1.96	9.65	73 (1)	50.19	1.22	9.66	64 (2)	24.39	2.45
e	10.04	34	26.61	1.92	10.08	49	45.35	1.57	10.09	83	21.82	2.26
f	—	—	—	—	10.52	11	44.54	1.37	—	—	—	—
12 a	7.46	28	42.18	1.39	7.53	15	101.67	0.80	7.67	22	30.58	1.92
b	8.55	53 (1)	41.04	1.26	8.48	43	80.15	1.25	8.57	58	25.13	1.92
c	9.18	44	28.60	1.67	9.21	53	56.98	1.08	9.25	62 (2)	23.31	2.21
d	9.64	44	28.51	1.79	9.63	61	46.13	1.26	9.63	68 (1)	22.12	2.28
e	10.05	32	21.92	1.96	10.06	40 (1)	38.04	1.76	10.05	55	17.40	3.05
f	—	—	—	—	—	—	—	—	—	—	—	—
13 a	7.62	18	56.48	0.80	7.56	18	83.20	0.86	7.55	23	54.80	1.37
b	8.56	69	31.62	1.50	8.49	50	62.67	1.23	8.48	72	31.34	1.67
c	9.22	55	27.25	2.10	9.24	51 (1)	49.55	1.76	9.20	50 (2)	18.06	2.58
d	9.64	43	21.79	2.26	9.63	69 (1)	48.00	1.26	9.67	60 (1)	21.07	2.26
e	10.06	42 (1)	22.29	2.21	10.09	58 (2)	37.98	1.68	10.05	44	16.89	2.56
f	—	—	—	—	10.59	16	29.05	2.07	—	—	—	—
14 a	7.65	15	40.80	1.20	7.63	12 (1)	107.39	0.36	7.53	18	55.16	1.23
b	8.51	51 (1)	39.78	1.57	8.53	27	68.46	0.80	8.46	49	34.16	1.68
c	9.21	48	33.88	1.37	9.24	34 (2)	57.98	1.08	9.18	44	33.25	1.50
d	9.66	48	28.23	1.67	9.69	62 (4)	55.77	1.36	9.66	46 (2)	18.66	2.56
e	10.09	48 (2)	20.14	2.26	10.11	72	40.62	1.32	10.09	58 (5)	18.44	2.56
f	10.60	22	24.47	2.03	10.72	43 (1)	39.96	1.32	—	—	—	—
15 a	7.57	10	78.36	0.64	7.56	10	123.56	0.54	7.53	18 (1)	58.14	1.08
b	8.54	39	44.53	1.26	8.50	33 (2)	73.35	0.86	8.50	30 (1)	47.12	1.08
c	9.23	32	38.21	1.32	9.23	40 (1)	59.25	0.97	9.19	37 (1)	23.61	1.96
d	9.62	32	34.88	1.37	9.67	51 (2)	53.70	1.08	9.65	42 (4)	27.53	2.07
e	10.07	31 (1)	23.14	1.96	10.08	61	53.28	1.08	10.09	55	22.46	2.56
f	10.63	18	20.47	2.64	10.61	72 (1)	44.05	1.57	10.62	38	24.39	2.19
16 a	7.59	14	59.71	1.36	7.38	11	174.4	1.25	7.42	20 (1)	44.57	1.08
b	8.48	35	61.06	0.80	8.58	25 (1)	81.92	0.80	8.40	36	42.36	1.39
c	9.20	23	41.76	1.37	9.24	40	76.40	0.80	9.10	26 (1)	36.68	1.44
d	9.67	30	34.10	1.37	9.68	42 (2)	69.38	1.08	9.67	35	22.47	1.96
e	10.11	28	32.70	1.37	10.11	78 (2)	57.66	1.08	10.10	55 (2)	29.60	1.67
f	10.59	18	31.33	1.67	10.67	74 (1)	51.12	1.20	10.71	55 (5)	19.35	2.56

5. *Estimate of Standard Error and Rejection of Observations.*—Perhaps the most difficult task in this work is to decide when an observation should be left out of the analysis, especially as the difficulty arises with exceptionally large values of μ , whose omission makes a big difference to the value of $\bar{\mu}$. Fig. 2 shows the distribution of the observed values of μ for a typical group, a distribution which is appreciably skew with a long tail in the positive direction.* It would be incorrect to apply the usual Chauvenet criterion because of the skewness. The method finally employed was somewhat arbitrary, as all decisions of this kind must be, but where there was room for doubt the observation was included in the analysis.

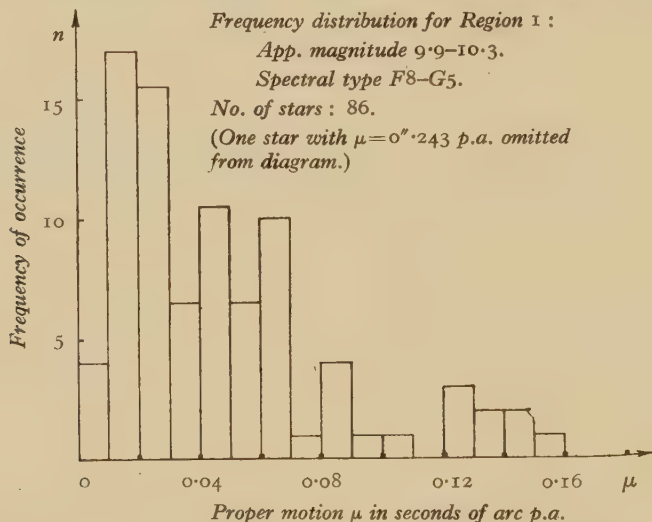


FIG. 2.—Frequency distribution of μ for group.

The method was as follows: from the observed values of μ the average, μ_1 , and the standard deviation, σ , were evaluated. It was found that μ_1 and σ were practically equal and that within the spectral divisions the values of σ were reasonably similar for all sixteen regions. It was decided as a first check to mark these stars for which

$$\mu > \mu_1 + 4\sigma \approx 5\mu_1$$

and then to examine the consequences of adopting this as a criterion for rejection. In doing so, μ_1 was available for each group and an estimate of σ with approximately 160 to 200 degrees of freedom was made by grouping the estimate from a number of regions. In most cases four representative regions were taken out of the sixteen, but for magnitude groups 6.9 to 7.9 and 10.4 to 11.2 eight were usually required. In one instance, namely A5 to F5 stars, σ was evaluated separately for each region for magnitude groups 6.9 to 7.9 and 10.4 to 11.2 because there was evidence in the former case at least of larger variations than usual in σ .

Having obtained a first approximation as above, any values close to the limiting value were examined in relation to their own region and to the general run of values for that magnitude group. From this it was possible to select a critical value μ_c for the group and any values of μ greater than this were neglected in the final solution.

* Taking $\log \mu$ instead of μ gives a more symmetrical frequency distribution, which is practically Gaussian in form. The author is examining this in connection with the evaluation of the correction c .

Table III shows the estimates of σ used for the first approximation and the limiting values μ_c finally chosen. (It will be seen that μ_c is approximately equal to 6σ .) The number of stars rejected by this procedure is shown in brackets after the number of stars used in Table II.

TABLE III

Distribution of Proper Motions : value of standard deviation of distribution σ and limiting value μ_c used as criterion in rejecting observations (observed values $> \mu_c$ were omitted in analysis).

[units : 0".001 p.a.]

App. Magnitude <i>m_{pg}</i>	Spectral Type					
	A5-F5		F8-G5		K0-M	
	σ	μ_c	σ	μ_c	σ	μ_c
6.9- 7.9	(60)*	(375)*	132.8	700	38.6	260
8.0- 8.9	29.8	190	65.4	320	23.5	160
9.0- 9.4	25.0	150	48.4	260	20.3	140
9.5- 9.8	18.3	120	41.6	250	18.8	125
9.9-10.3	16.3	120	40.1	230	21.7	135
10.4-11.2	(22)*	(125)*	38.9	210	18.2	125

[* In these cases the value of σ varied appreciably from region to region. The values in brackets represent average values for σ and the largest value of μ included in the analysis. All the stars in these groups were included.]

Finally an estimate of the uncertainty attached to each value of the mean parallax $\bar{\pi}$ is required and here two different estimates are possible. The first would be that corresponding to repeated observations of the same group of stars. Since there is a probable error attached to the observed value of μ_α and μ_δ and to the parameters in the two-stream theory estimated for each region, repeating the observations on the same stars would not yield exactly the same value of $\bar{\pi}$. Thus we could consider a frequency distribution of $\bar{\pi}$, corresponding to repeated determinations from the same stars, with an appropriate standard deviation. We are more interested, however, in the standard deviation arising from the variations in proper motion among the stars considered. If we regard them as if they were observations of a single variable whose average value we wish to estimate, then this average value will have a standard error

$$\sigma(\bar{\mu}) = \text{standard error of } \bar{\mu} = \sigma/\sqrt{N},$$

where N is the number of stars in the group. (Although the distribution of μ is skew the derived distribution of $\bar{\mu}$, the average value of μ for a group of 30 to 60 stars, will be adequately described by its standard error.)

We should also take into account the uncertainty attached to D in equation (1), but an examination of the magnitude of the latter showed that it would increase the estimate obtained on neglecting it by less than 5 per cent. (much less in most cases). For this reason the formula used was

$$\sigma_1 = \text{standard error of } \bar{\pi} = \frac{0.242\sigma}{DN^{\frac{1}{2}}}. \tag{3}$$

Values of Mean Parallax ($\bar{\pi}$, unit $0''.001$), Standard Error of Estimate (σ_1), and Number of Stars used (N)

Region	Galactic latitude b	$m=6.9$ to 7.9			$m=8.0$ to 8.9			$m=9.0$ to 9.4		
		$\bar{\pi}$	σ_1	N	$\bar{\pi}$	σ_1	N	$\bar{\pi}$	σ_1	N

(a) A5-F5

1	-72.2	9.10	1.80	12	8.84	0.88	33	7.03	0.80	28
2	-64.2	8.40	1.12	10	7.98	0.77	44	5.62	0.84	26
3	-49.7	7.51	1.20	9	7.13	0.66	56	5.16	0.78	29
4	-34.0	10.02	2.70	18	7.83	0.79	51	5.63	0.68	48
5	-18.8	10.45	2.30	18	4.22	0.51	83	4.53	0.46	74
6	-5.1	8.07	1.34	23	5.02	0.55	71	3.10	0.50	61
7	-6.2	5.28	0.60	29	4.23	0.43	72	3.48	0.36	70
8	$+13.7$	11.48	1.98	30	5.30	0.57	52	4.79	0.46	56
9	$+16.2$	8.58	1.86	18	4.76	0.48	63	3.93	0.44	50
10	$+13.1$	12.46	2.31	15	4.84	0.55	54	4.10	0.46	54
11	$+5.1$	6.36	1.11	21	4.49	0.54	47	3.39	0.38	63
12	-6.5	6.21	0.86	28	6.04	0.60	53	4.21	0.56	44
13	-20.5	8.93	1.26	18	5.00	0.56	69	4.31	0.53	55
14	-35.8	6.95	1.08	15	6.77	0.71	51	5.77	0.62	48
15	-51.4	14.51	3.17	10	8.25	0.88	39	7.07	0.82	32
16	-65.5	10.75	3.04	14	10.99	0.90	35	7.52	0.94	23

(b) F8-G5

1	-72.2	32.07	7.89	9	16.40	1.98	35	10.87	1.50	33
2	-64.2	17.97	5.65	17	11.71	1.96	34	9.75	1.46	34
3	-49.7	16.18	5.65	18	12.52	1.84	41	10.38	1.28	46
4	-34.0	20.67	9.01	7	12.19	1.86	40	10.58	1.28	46
5	-18.8	22.95	6.37	14	9.68	1.81	42	10.66	1.26	47
6	-5.1	14.81	5.83	15	8.95	1.46	58	7.70	1.02	65
7	$+6.2$	10.77	5.25	16	9.04	1.50	48	7.21	0.94	66
8	$+13.7$	10.53	6.04	11	11.15	1.37	52	8.36	1.08	46
9	$+16.2$	23.09	4.95	13	10.27	1.60	30	7.27	0.94	48
10	$+13.1$	16.88	5.52	14	11.84	1.47	48	8.17	1.02	54
11	$+5.1$	9.93	4.37	18	9.84	1.34	46	7.24	0.90	57
12	-6.5	14.42	5.85	15	11.37	1.42	43	8.08	0.94	53
13	-20.5	14.03	5.26	18	10.57	1.56	50	8.36	1.14	51
14	-35.8	19.72	7.04	12	12.57	2.31	27	10.65	1.52	34
15	-51.4	21.67	7.36	10	12.86	2.00	33	10.39	1.34	40
16	-65.5	32.59	7.55	11	15.31	2.44	25	14.28	1.43	40

(c) Ko-M

1	-72.2	11.64	2.26	12	7.41	0.94	26	4.94	0.76	29
2	-64.2	7.54	2.41	8	4.73	0.68	37	5.10	0.70	26
3	-49.7	8.99	2.14	12	3.91	0.75	36	4.77	0.67	34
4	-34.0	5.22	2.16	12	4.76	0.74	38	5.12	0.67	34
5	-18.8	10.45	2.05	13	5.21	0.55	67	3.24	0.50	59
6	-5.1	4.99	1.08	36	3.50	0.42	89	3.35	0.40	56
7	-6.2	7.48	1.04	31	3.95	0.44	65	3.42	0.36	69
8	$+13.7$	5.38	1.27	21	4.50	0.40	78	4.16	0.32	89
9	$+16.2$	7.37	1.06	29	4.42	0.47	54	3.07	0.40	56
10	$+13.1$	7.27	1.20	24	5.28	0.45	64	3.34	0.37	69
11	$+5.1$	7.18	1.14	31	4.57	0.52	54	3.74	0.40	70
12	-6.5	4.31	1.16	22	3.54	0.44	58	3.29	0.36	62
13	-20.5	9.31	1.36	23	5.34	0.47	72	3.07	0.49	50
14	-35.8	8.86	1.46	18	5.49	0.54	49	5.34	0.49	44
15	-51.4	11.61	1.82	18	9.41	0.86	30	4.71	0.66	37
16	-65.5	7.79	1.51	20	7.41	0.68	36	6.41	0.70	26

TABLE IV (cont.)

Region	Galactic latitude	$m=9.5$ to 9.8			$m=9.9$ to 10.3			$m=10.4$ to 11.2		
	b	$\bar{\pi}$	σ_1	N	$\bar{\pi}$	σ_1	N	$\bar{\pi}$	σ_1	N
(a) A5-F5										
1	-72.2	5.47	0.66	22	5.57	0.52	29			
2	-64.2	7.67	0.70	20	5.05	0.64	19			
3	-49.7	4.38	0.59	27	4.39	0.55	24			
4	-34.0	5.01	0.54	41	4.58	0.55	32			
5	-18.8	3.97	0.36	60	2.80	0.46	30			
6	-5.1	2.99	0.44	43	3.48	0.57	20			
7	$+6.2$	3.18	0.30	56	3.12	0.42	22			
8	$+13.7$	3.69	0.41	38	3.35	0.50	20			
9	$+16.2$	2.77	0.34	47	3.10	0.36	33			
10	$+13.1$	3.70	0.33	56	3.16	0.32	49			
11	$+5.1$	2.80	0.32	50	3.27	0.34	34			
12	-6.5	4.19	0.40	44	3.22	0.42	32			
13	-20.5	3.44	0.44	43	3.52	0.40	42			
14	-35.8	4.81	0.45	48	3.42	0.40	48	4.17	0.64	22
15	-51.4	6.46	0.60	32	4.28	0.54	31	3.79	0.72	18
16	-65.5	6.14	0.60	30	5.89	0.56	28	5.64	1.26	18
(b) F8-G5										
1	-72.2	11.18	1.00	56	8.40	0.84	85	8.38	1.07	42
2	-64.2	9.84	1.01	52	7.44	0.90	60	8.28	1.20	32
3	-49.7	9.23	0.87	75	8.05	0.86	70	6.37	1.66	18
4	-34.0	6.94	0.88	73	6.57	0.89	69	6.04	2.47	8
5	-18.8	5.52	0.96	60	8.95	1.32	30	—	—	—
6	-5.1	6.47	1.15	38	4.64	1.96	12	—	—	—
7	$+6.2$	7.12	1.12	35	5.58	1.27	25	—	—	—
8	$+13.7$	6.78	0.96	43	7.90	1.05	33	7.47	1.86	10
9	$+16.2$	6.98	0.59	90	6.23	0.66	79	5.56	1.30	16
10	$+13.1$	6.06	0.77	71	6.79	0.70	79	4.34	1.91	10
11	$+5.1$	7.00	0.68	73	6.33	0.80	49	6.21	1.64	11
12	-6.5	6.54	0.76	61	5.40	0.90	40	—	—	—
13	-20.5	8.09	0.84	69	6.40	0.89	58	4.90	1.64	16
14	-35.8	10.24	0.97	62	7.46	0.87	72	7.34	1.09	43
15	-51.4	9.42	1.02	51	9.34	0.90	61	7.72	0.80	72
16	-65.5	12.97	1.20	42	10.78	0.85	78	9.55	0.84	74
(c) Ko-M										
1	-72.2	4.96	0.68	32	5.27	0.70	40	4.96	0.66	32
2	-64.2	3.35	0.88	14	4.44	0.65	35	4.12	0.72	20
3	-49.7	4.05	0.64	31	4.06	0.60	48	3.77	0.73	23
4	-34.0	4.08	0.48	59	3.81	0.60	49	3.52	0.82	19
5	-18.8	3.77	0.40	78	3.26	0.86	23	—	—	—
6	-5.1	2.79	0.41	60	4.58	1.01	13	—	—	—
7	$+6.2$	3.05	0.37	58	3.45	0.40	64	—	—	—
8	$+13.7$	3.77	0.36	60	2.77	0.40	66	—	—	—
9	$+16.2$	3.69	0.33	69	2.64	0.33	94	2.24	0.48	31
10	$+13.1$	4.18	0.40	52	3.28	0.36	81	—	—	—
11	$+5.1$	4.02	0.38	64	3.59	0.39	83	—	—	—
12	-6.5	3.12	0.32	68	2.45	0.41	55	—	—	—
13	-20.5	3.58	0.41	60	2.87	0.56	44	—	—	—
14	-35.8	3.00	0.44	46	2.96	0.46	58	—	—	—
15	-51.4	5.50	0.58	42	4.48	0.58	55	4.87	0.59	38
16	-65.5	3.93	0.56	35	5.18	0.51	55	3.38	0.43	55

6. *Values of Mean Parallax for each Region.*—Table IV shows the values of $\bar{\pi}$ for each group of stars, *i.e.* sub-divided according to spectral type and apparent magnitude, together with the number of stars in the group (N) and the standard error of the estimate (σ_1). The outstanding features are the increase of $\bar{\pi}$ with galactic latitude and the difference between F8 to G5 stars and the other two spectral divisions. The most reliable results are for magnitude groups 8.0 to 8.9, 9.0 to 9.4 and 9.5 to 9.8, where the standard error is of the order of 10 per cent. of the mean parallax.

7. *Grouping of Results and Comparison with Groningen Values.*—To get a more reliable comparison with the Groningen values the regions were grouped according to galactic latitude as shown in Table V, which also gives the mean galactic latitude of each group taken without regard to sign. (The data did not seem sufficient to discriminate between the northern and southern hemispheres.) The results for the combined regions are shown in Table VI; in this table

$$\begin{aligned}\pi &= \text{mean parallax for group} \\ &= \Sigma N \bar{\pi} / \Sigma N,\end{aligned}\quad (4)$$

where N = number of stars in each region, and the summation is over the regions grouped together,

$$\begin{aligned}s &= \text{standard error of estimate of } \pi \\ &= \{\Sigma (N \sigma_1^2)\}^{\frac{1}{2}} / \{\Sigma N\}.\end{aligned}\quad (5)$$

$$N' = \text{number of stars in grouped regions} = \Sigma N. \quad (6)$$

$$m = \text{mean apparent photographic magnitude.}$$

The value of m has been rounded off to 0^m.05.

TABLE V
Grouping of Regions for Summary

Identification Letter	Regions	Mean Galactic Latitude (neglecting sign)	Magnitude Correction δm
A	1, 2, 3, 15, 16	60	-0.15
B	4, 5, 13, 14	27	-0.15
C	8, 9, 10	14	0.0
D	6, 7, 11, 12	6	0.0
E	14, 15, 16	51	-0.1
F	4, 13, 14	32	-0.2
G	4 to 14 inclusive.	16	0.0

Roughly, $s = \pi / (N')^{\frac{1}{2}}$, and it is possible to guess the accuracy of the estimate of π from the number of stars used in the analysis.* For the three best-determined magnitude groups, between magnitude 8.0 and 9.8, s is approximately 7 per cent. of π , an accuracy comparable with the Groningen determinations.

For a comparison with the latter we must reduce the results to the Harvard visual scale m_v , for which the Groningen values are tabulated. This introduces two sources of error,

- (i) the Cape magnitude scale varies with right ascension,
- (ii) the colour index will not be the same for the spectral classes included in our groups.

* Conversely this approximate relationship could be used to decide the number of stars which need be included to determine π to a given accuracy.

TABLE VI

Summary of Results: Variation of Mean Parallax with Galactic Latitude and Spectral Type
[unit for π and s is $0''.001$]

Group	A5-F5				F8-G5				Ko-M			
	π	s	N'	m	π	s	N'	m	π	s	N'	m
A	10.1	1.1	55	7.55	22.5	2.95	65	7.55	9.6	0.85	70	7.45
A	8.4	0.35	207	8.50	13.7	0.90	168	8.55	6.4	0.35	165	8.50
A	6.5	0.35	138	9.20	11.2	0.60	193	9.25	5.1	0.30	152	9.20
A	6.0	0.3	131	9.65	10.3	0.45	276	9.65	4.5	0.30	154	9.65
A	5.0	0.25	131	10.10	8.8	0.40	354	10.10	4.7	0.25	233	10.10
E	4.5	0.5	58	10.60	—	—	—	—	—	—	—	—
A	—	—	—	—	8.4	0.45	238	10.60	4.2	0.25	168	10.65
B	9.2	1.0	69	7.60	18.7	3.3	51	7.55	8.7	0.85	66	7.50
B	5.7	0.3	254	8.55	11.1	0.9	159	8.55	5.2	0.3	226	8.50
B	5.0	0.3	225	9.20	10.0	0.65	178	9.20	4.1	0.25	197	9.20
B	4.3	0.2	192	9.65	7.7	0.45	264	9.65	3.6	0.2	243	9.65
B	3.6	0.2	152	10.05	7.1	0.50	229	10.05	3.2	0.3	174	10.05
F	—	—	—	—	6.6	0.85	67	10.70	—	—	—	—
C	10.9	1.2	63	7.55	17.2	3.15	38	7.55	6.8	0.65	74	7.55
C	5.0	0.3	169	8.55	11.2	0.85	130	8.55	4.7	0.25	196	8.55
C	4.3	0.25	160	9.20	7.9	0.6	148	9.25	3.6	0.2	214	9.20
C	3.4	0.2	141	9.65	6.6	0.4	204	9.65	3.9	0.2	181	9.65
C	3.2	0.2	102	10.05	6.8	0.45	191	10.05	2.9	0.2	241	10.05
C	—	—	—	—	5.8	0.95	36	10.50	—	—	—	—
D	6.4	0.5	101	7.50	12.3	2.65	64	7.60	6.1	0.55	120	7.55
D	4.9	0.25	243	8.55	9.7	0.7	195	8.50	3.8	0.25	266	8.55
D	3.5	0.2	238	9.20	7.5	0.5	241	9.20	3.4	0.2	276	9.25
D	3.3	0.2	193	9.65	6.8	0.45	207	9.65	3.2	0.2	250	9.65
D	3.3	0.2	108	10.05	5.7	0.55	126	10.05	3.3	0.25	215	10.05
G	8.4	0.5	233	7.55	15.7	1.75	153	7.55	6.9	0.4	260	7.55
G	5.2	0.15	666	8.55	10.6	0.5	484	8.55	4.6	0.15	688	8.55
G	4.2	0.15	623	9.20	8.4	0.3	567	9.20	3.7	0.15	687	9.20
G	3.7	0.1	526	9.65	7.1	0.25	675	9.65	3.6	0.1	674	9.65
G	3.4	0.15	362	10.05	6.7	0.3	546	10.05	3.1	0.15	630	10.05

The Cape magnitude scale was discussed in the original volume* which gave the spectral types and magnitudes of the stars under consideration and again in a catalogue of faint stars in the Cape Astrographic Zone. Table I† of the introduction to the latter gives a comparison of the Cape magnitudes with other magnitude scales for the Kapteyn Areas contained in the zone. Corrections δm for the regions of the present paper were taken from column 7 of this table, with occasional references to column 5 when in doubt. This gives a correction to reduce Cape photographic magnitudes to Harvard photographic magnitudes.

* Cape Catalogue 1, Introduction.

† Cape Catalogue 3, ix.

It was assumed that the error was a zero error only and that there was no colour effect for different spectral types.* The values taken for δm , in the sense

$$\text{Harvard} = \text{Cape} + \delta m,$$

are shown in the last column of Table I and for the grouped regions in the last column of Table V. For the colour indices, to transfer from photographic magnitudes to visual magnitudes, the values used were

Spectral group	A5-F5	F8-G5	K0-M
$m_p - m_v$	0.4	0.7	1.4

As Van Rhijn and Bok§ give secular parallaxes for different galactic latitudes and visual magnitudes, these values are not equal to the mean parallax π of Table VI but proportional to it, *i. e.*

$$(\overline{h/\rho}) = \text{Groningen secular parallax}$$

$$= V_0 \pi / \kappa,$$

where

$$V_0 = \text{solar motion in km./sec.,}$$

and 1 Astronomical Unit/year = κ km./sec. = 4.74 km./sec.

Taking V_0 as 19.5 km./sec. as before,

$$\pi = 0.243(\overline{h/\rho}). \quad (7)$$

Table VII repeats the final Groningen table† in terms of π .

TABLE VII

Groningen Publ. No. 45: Values of Mean Parallax for different Visual Magnitudes and Zones of Galactic Latitude

[Tabulated in units 0".001, using formula mean parallax = 0.243 (secular parallax)]

Visual Magnitude, m_v	Galactic Latitudes and Mean Galactic Latitude			
	0° to ±20°	±20° to ±40°	±40° to ±90°	0 to ±90°
	10°	30°	57°	
5.0	14.3	17.0	19.9	16.5
6.0	10.4	11.7	15.1	11.9
7.0	7.8	8.0	11.4	8.8
8.0	5.6	5.6	8.8	6.3
9.0	4.1	4.1	6.6	4.6
10.0	2.9	2.9	4.9	3.2
11.0	2.1	2.1	3.6	2.4
12.0	1.5	1.5	2.9	1.7

To assist in comparing the values of π , Figs. 3, 4, 5 and 6 were drawn from the data of Tables VI and VII. In these diagrams the difference between the spectral divisions is immediately obvious, the Groningen values occupying an intermediate position close to the values for A5-F5 stars. In Fig. 3, for the low galactic latitude regions, where the stars are most numerous, the plotted points for K0-M stars and for A5-F5 stars suggests that the star of apparent magnitude 9.9 to 10.3 were not selected at random, but that there was a tendency to include too many stars of large proper motion. As the limiting magnitude of the catalogue is likely to be lowest for these regions this tendency seems plausible and the effect is small in any case.

* Cf. *Bull. astr. Obs. Harv.*, 901, pp. 25-30, 1935.

† *Loc. cit.*, p. 21, Table 8.

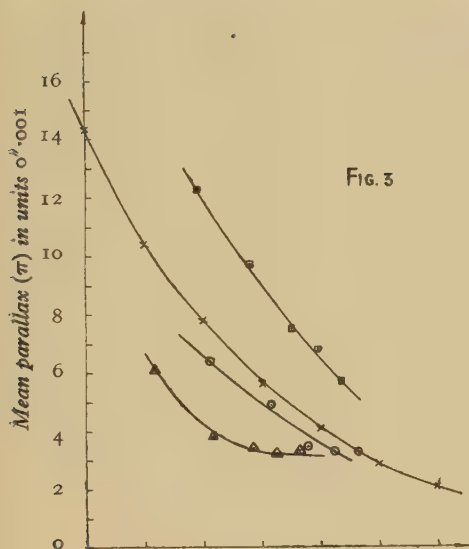


FIG. 3

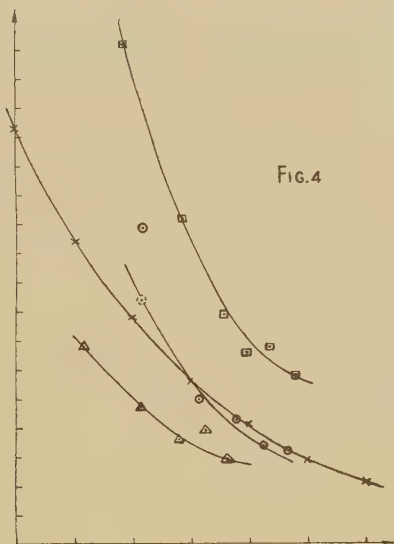


FIG. 4

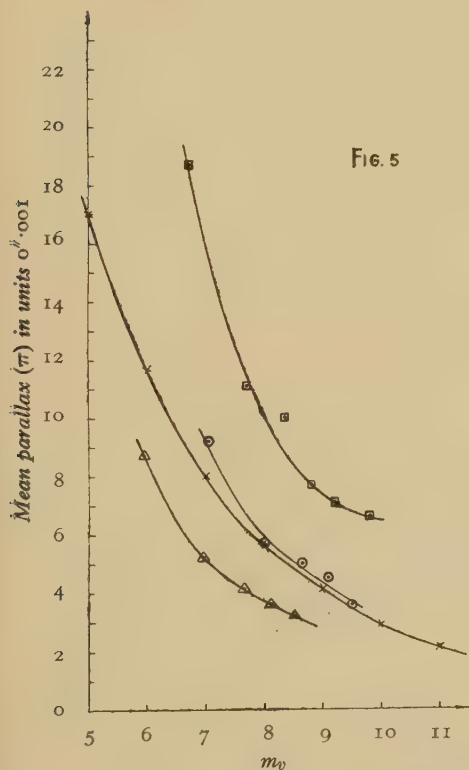


FIG. 5

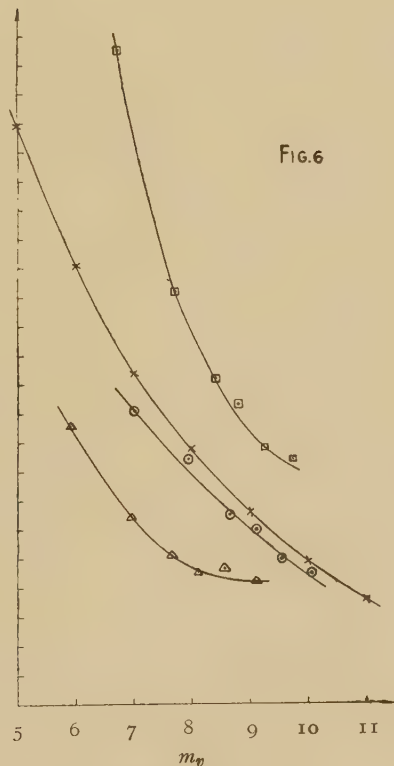


FIG. 6

Graphs of mean parallax against apparent magnitude for different spectral types.

Key: Crosses indicate Groningen values.

Squares indicate Cape Stars, F8-G5.

Circles indicate Cape Stars, A5-F5.

Triangles indicate Cape Stars, K0-M.

FIG. 3.—Low galactic latitudes: Groningen values $b=10^\circ$, Cape stars $b=6^\circ$.

FIG. 4.—Low galactic latitudes: Groningen values $b=10^\circ$, Cape stars $b=14^\circ$.

FIG. 5.—Medium galactic latitudes: Groningen values $b=30^\circ$, Cape stars $b=27^\circ$.

FIG. 6.—High galactic latitudes: Groningen values $b=57^\circ$, Cape stars $b=60^\circ$.

For low galactic latitudes Fig. 3 shows the values found for regions 6, 7, 11 and 12, whose centres lie 6° on either side of the galactic equator, while Fig. 4 shows the values for regions 8, 9 and 10, a group of regions in the northern galactic hemisphere. In the latter case a check can be obtained by combining regions 4 to 14 inclusive, since the mean numerical value of the galactic latitude for these regions is close to 14° . It will be seen from Table VI that the difference is negligible except for A5-F5 stars of apparent magnitude 6.9 to 7.9. In drawing Fig. 4, the values from regions 8, 9 and 10 were used, save for the exceptional instance above where the values from regions 5 to 14 was plotted.* The Groningen values in this figure are the same as in Fig. 3.

8. *Variation of Mean Parallax with Galactic Latitude.*—As the present values show the increase of parallax with galactic latitude more clearly than the Groningen values do, the parallaxes for the separate spectral types were combined to give the graphs in Fig. 7. This was done rather crudely by reading off the values of π for

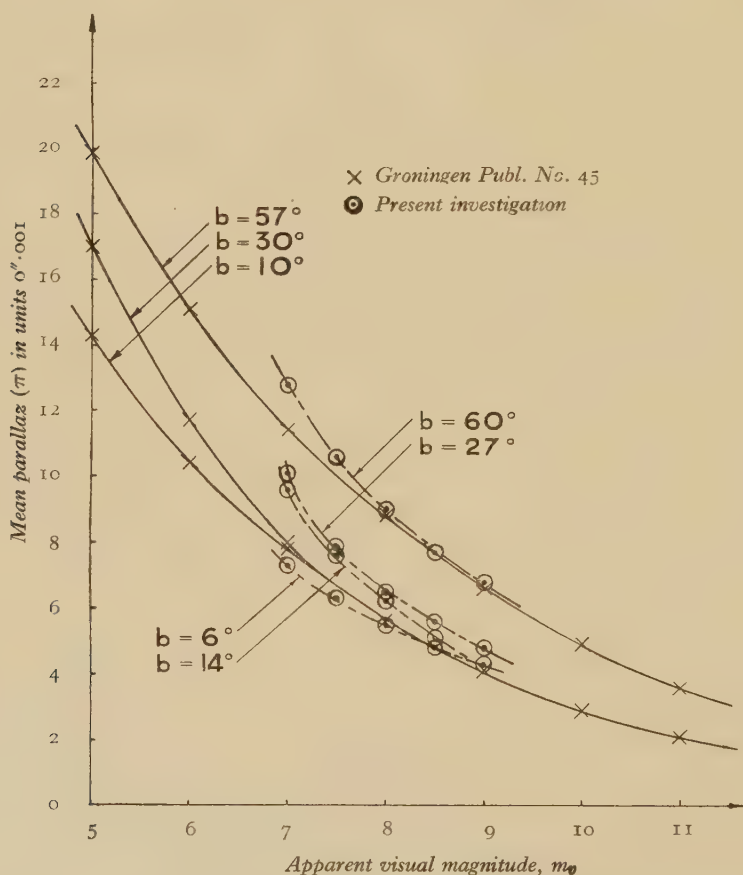


FIG. 7.—Variation of parallax with galactic latitude.

$m = 7.0, 7.5, 8.0, 8.5$ and 9.0 from the curves drawn through the plotted points in Figs. 3 to 6, and then weighting these values according to the number of stars in each spectral division for all magnitudes. The Groningen values are repeated for

* In the diagram this point is shown with a broken circle to distinguish it.

comparison. The tendency is for the values from the present investigation to be higher than those of Groningen, especially for the brighter stars. The difference between the two sets of values in medium galactic latitudes is accounted for by the way in which the Groningen values for low and medium galactic latitudes were grouped in forming the final table (*cf.* Table VII).

9. *Discussion of the Results.*—Although there are a number of possible sources of error or factors which might affect the results, the only one of real importance is the assumption about the frequency distribution of velocities implicit in the formula for the mean transverse velocity. In other cases, either an allowance could be made for the effect or it is likely to be small. For example, the solar velocity has been taken as 19.5 km./sec. for all three spectral divisions; this does not affect the comparison with Groningen values since the same solar velocity was used in forming Table VII, but it could be allowed for by multiplying the figures in Table VI by a suitable ratio. If we take Campbell and Moore's values* as giving approximately

Spectral group	A5-F5	F8-G5	K0-M
Solar velocity	19.5	18.5	20.0 km./sec.

then we should have to multiply the values for F8-G5 stars by $19.5/18.5 = 1.05$, and to multiply the values for K0-M stars by $19.5/20.0 = 0.97$. In either case the difference is less than the standard error of π and it would be over-estimating the accuracy of the results to include the correction.†

Similarly an allowance could be made for the effects of galactic absorption, but as the Groningen values do not appear to have been corrected for absorption it might be better to follow their example and regard the "apparent magnitude" as including the effect of absorption. The error involved in taking

$$\text{mean parallax} = (\text{mean proper motion})/(\text{mean transverse motion})$$

instead of

$$\text{mean parallax} = \text{mean value of (proper motion/transverse motion)}$$

is unknown. The magnitude groups and spectral groupings have been kept as small as possible to make the groups homogeneous, since the error depends on the distribution of distance and transverse velocity for the stars concerned. An examination of the distribution of μ may throw some light on the matter, but it seems unlikely that the error will be appreciable compared with the error in the mean transverse velocity which would be occasioned by departures from the assumed frequency distribution of velocities.

At first sight the results appear to confirm the hypothesis on which the calculations were based, since the final values for all spectral types agree well with the Groningen values. The slightly more rapid increase of mean parallax with galactic latitude is reasonable; if anything it is a point in favour of the present results. Further, the Groningen values are based mainly on stars in the northern

* *Publ. Lick Obs.*, **16**, 1928. These values are for stars brighter than 5.51 visual magnitude and include stars in N. and S. hemispheres.

† It is possible that for fainter stars the solar velocity is larger than Campbell and Moore's figures (*cf. Handb. Astrophys.* VII, 600, 1936), but the evidence on this is not conclusive. For G-type stars, Nordström's values give $V_0 = 14$ km./sec. for giants and $V_0 = 31$ km./sec. for dwarfs; for an equal number of giants and dwarfs (*cf. Van Rhijn and Schwassman, Z. Astrophys.* **10**, 161, 1935) this brings us back to values near 20 km./sec.

hemisphere, and asymmetry between the northern and southern hemispheres would not be surprising.* So at first sight the agreement is good. Against that, the number of stars rejected is large among F8-G5 and K0-M stars, namely 61 out of 3895 and 84 out of 4005 respectively. Although this question involves the distribution of the stars in space as well as the frequency distribution of the stars' velocities, it recalls the suggestion made by a number of authors that neither the two-stream theory nor the ellipsoidal frequency distribution predicts sufficient stars with high velocities.

Comparing the results with mean parallaxes for the same stars which Jackson gives in his introduction to the Catalogue†, there are striking differences in the values obtained. Jackson's solution was obtained as part of a determination of the solar motion by Airy's method, which takes no account of preferential motion. In solving, results for the whole zone were considered together irrespective of galactic latitude, so the figures in Table VI have been grouped in a similar manner for comparison. The comparison is shown in Table VIII.

TABLE VIII
*Comparison of Results with Jackson's Mean Parallaxes
for the same Stars, grouping all Regions together*
[The table gives mean parallaxes in units $0''.001$]

Jackson's Values				Present Investigation			
Magnitude Range, m_{pg}	Spectral Type			Magnitude Range, m_{pg}	Spectral Type		
	A5-F5	F8-G5	K0-M		A5-F5	F8-G5	K0-M
≤ 7.9	9.4	12.1	7.3	6.9-7.9	8.8	17.7	7.5
8.0-8.9	6.8	10.3	5.4	8.0-8.9	6.0	11.4	4.9
9.0-9.9	4.9	7.7	4.2	9.0-9.4	4.6	9.1	4.0
				9.5-9.8	4.1	8.0	3.8
≥ 10.0	3.6	6.2	3.0	9.9-10.3	3.8	7.4	3.6
				10.4-11.2	—	7.7	3.8

The biggest difference is for F8-G5 stars, where the values of the present paper are greater than Jackson's values; for A5-F5 stars the trend is in the opposite direction, although the differences are small compared with the standard errors of the determinations, while for K0-M stars the results are about the same. In his solution, Jackson omitted all stars for which the proper motion in either component exceeded $0''.2$ per annum, whereas Table III shows that this is much too drastic for F8-G5 stars. Reference to Table II‡ of Jackson's introduction to the catalogue gives 158 stars of type F8-G5 which would be omitted, roughly one star in 25.

For high galactic latitudes, Van Rhijn and Schwassman§ evaluated mean parallaxes for stars of visual magnitude 9.0 for spectral types F, G, K and M from

* Cf. the difference in the number of stars per square degree between the Cape catalogues (Cape Catalogue 3, xvi) and *Publ. astr. Lab. Groningen*, No. 43, 1929.

† Cape Catalogue 2, xxviii-xxx.

‡ Cape Catalogue 2, xxlii-xxliv.

§ *Z. Astrophys.*, 10, 161-187, 1935.

proper motions given in the *Yale Catalogue* for declination $+50^\circ$ to $+55^\circ$. In deriving them, they used the same formula as Van Rhijn and Bok in *Publ. astr. Lab. Groningen*, No. 45,

$$\text{mean parallax} = (C \Sigma v \sin \lambda) / (\Sigma \sin^2 \lambda),$$

where

v = component of parallactic motion towards antapex,

λ = angular distance between star and solar antapex,

and C is a constant depending on the Solar Motion.

Although the results of Van Rhijn and Schwassman* are for a mean galactic latitude of 50° and are sub-divided differently, it can be said from inspection that the agreement for early and late spectral types is reasonable, but for F and G stars the present investigation gives larger mean parallaxes. The agreement for K stars confirms their conclusion that even in high galactic latitudes 90 per cent. of the K-type stars with apparent (photographic) magnitudes less than 9.5 are giants with absolute magnitudes less than $+3.0$.

Acknowledgments.—The work published in this paper was started in the summer of 1940 and resumed after an interval in 1945. I should like to take this opportunity of thanking Professor W. M. Smart for suggesting the investigation, for supervising it in the early stages, and for a number of helpful discussions on the subject since. I would acknowledge also the help of Dr Tannahill in putting at my disposal unpublished tables and preliminary working from the analysis of the directions of the proper motions. Finally, I am indebted to Dr Brück for advice and encouragement while working at Cambridge University Observatory.

Trinity College,
Cambridge:

1946 March 25.

* *Loc. cit.*, p. 177, Table 10.

THE GROWTH AND DECAY OF THE SUNSPOT MAGNETIC FIELD

T. G. Cowling

(Received 1946 February 2)

Summary

The variation of the magnetic field at the centre of a sunspot is studied for spots of lifetime 30 and 55 days. The study shows that, if the magnetic field is produced by the darkening of the spot, a time at most of order one day is taken to establish the field. Theory indicates that 300 years is required for the growth or decay of the field. It is concluded that the sunspot field must exist both before and after the visible lifetime of the spot.

1. An adequate theory of a sunspot must explain the coolness of the spot, its magnetic field, and its dynamical properties. The causal relation between these is obscure. Most theoretical workers seek to explain the coolness by the dynamical properties, and the magnetic field by the coolness; students of magnetism, on the other hand, often seek to make the magnetic field fundamental.

The present paper describes a study of the magnetic history of large sunspots. The results indicate fairly clearly that the magnetic field cannot grow and decay *in situ*; it must exist before the spot becomes visible, and persist after the spot disappears. Since, however, a strong magnetic field is seldom observed in the absence of a visible spot, the field existing before and after the spot must either be below the visible layers, or be very different in form from that visible in a spot. Whichever of these is correct, the magnetic field can clearly not be ignored in considering the origin of a spot.

2. The growth and decay of the field of large spots was studied from the diagrams given in the Mount Wilson volume *Magnetic Observations of Sunspots*, 1916-1924. The results illustrated in this remarkable volume are, in one respect, difficult to interpret, because of a "foreshortening" effect, whereby a spot shows a smaller magnetic field when on the Sun's limb than when near the centre of the disk. Three possible reasons for this may be advanced. Geometrical foreshortening makes the identification of small patches of intense field more difficult; the large inclination of the field to the line of sight near the limb causes the extinction of the wing lines of a Zeeman triplet by a polarizing apparatus to be only partial; and if a sunspot is lower than the Sun's surface around it, as found last century by Wilson and Carrington, and recently confirmed by Pettit*, the regions of intense field in a sunspot at the limb may be hidden by obscuring clouds. The magnetic evidence itself suggests that a sunspot is a depression in the Sun's surface. When a spot is near the limb, the far part exhibits a magnetic polarity opposite in sign from the rest; and often this part appears to include most of the area of the spot, though it should actually be a small part. This indicates that the near part of normal polarity is obscured by higher material outside the spot.

The correction for magnetic "foreshortening" was obtained from the middle part of the history of long-lived spots. The intensity of the field of a spot visible

* *Mount Wilson Reports*, p. 12, 1939.

during parts of four or more revolutions varies very little during the middle passages of the spot across the Sun's disk ("middle" passages are understood to be all save the first and last). Accordingly the passages were studied for spots visible during parts of *three* or more revolutions. The maximum observed intensity in each spot (*i. e.* the central intensity) was tabulated for each of twelve zones of solar longitude, 15° in breadth, from 90° W. to 90° E. Strictly the variation with angular distance from the centre of the Sun's disk should have been considered, rather than that with longitude; but the longitude calculations appear adequate and were much simpler.

In the tabulation, certain spots were rejected, because of serious incompleteness in the observational record, or suspected intermittency of appearance. In all, 45 passages of spots across the Sun's disk were tabulated. The observed field for each spot was multiplied by a suitable factor, to make the maximum field observed during the passage about 3000 gauss; this rarely involved a change of more than 10 per cent. The figures so obtained were averaged, giving the following table. Little reliance can be placed on the fields shown for the zones 90°-75°.

Longitude	Number of spots averaged	Field* (gauss)	2000 + 900 cos ϕ
90°-75° E.	17	2000?	2100
75°-60°	33	2300	2350
60°-45°	33	2550	2550
45°-30°	37	2700	2700
30°-15°	38	2850	2825
15°- 0°	40	2850	2900
0°-15° W.	39	2900	2900
15°-30°	41	2800	2825
30°-45°	43	2700	2700
45°-60°	41	2600	2550
60°-75°	30	2400	2350
75°-90°	13	2300?	2100

The fields given in this table show no appreciable E.-W. asymmetry. As the figures given in the last column indicate, they are closely represented by the expression $2000 + 900 \cos \phi$, where ϕ is the longitude. Thus, to correct observed fields for "foreshortening", the latter must be multiplied by $29/(20 + 9 \cos \phi)$. In passing, we may note that similar calculations, using firstly spots visible during only three revolutions, and then spots visible during more than three revolutions, gave results not significantly different.

3. The correction factor was used in studying spots of lifetime about 30 days. Again spots inadequately observed, or suspected of intermittency, were rejected: The start of their lifetime was studied from spots visible during two successive revolutions, and born on the disk (38 spots); the maximum field observed in a spot was averaged for the first and each successive day of the spot. The final part was similarly studied from spots visible during two successive revolutions, and dying out on the disk (39 spots); the field was averaged for the last and each preceding day of the spot. The variation of the field during the middle of the lifetime has to be inferred from the behaviour of such large spots as are visible right across the Sun's disk during one revolution and not otherwise seen, and show no sign of the immediate proximity of birth or death while visible (44 spots).

These indicate that, during the middle of the spot's life, the field is nearly constant and equal to about 2700 gauss; the error of this estimate is not likely to be much more than 100 gauss.

A similar calculation was made for spots of lifetime about 55 days, for which, however, the variation of the field during the middle part of the lifetime was less uncertain. The results of the calculations are shown in Fig. 1, in which the Greenwich curves showing the variation in area of 30-day and 60-day spots are also shown.* A remarkable feature of the magnetic curves is their flatness near maximum. The curve for 55-day spots indicates a maximum field not much greater than that for the 30-day spots. In fact, for very few spots does the field greatly exceed 3000 gauss, save spasmodically and for short times.

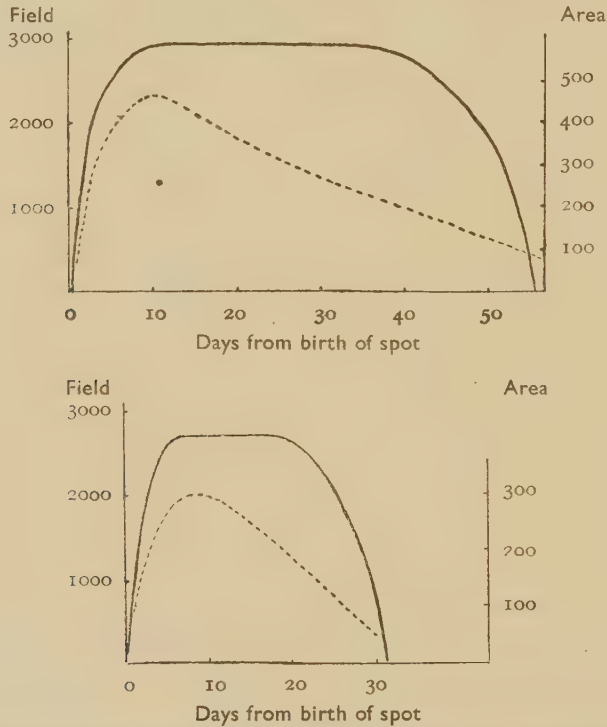


FIG. 1.—Central field (continuous curves) in gauss, and areas (broken curves) in millionths of the Sun's visible hemisphere, plotted against days from birth of spot. Upper curves, field for 55-day spots, area for 60-day spots; lower curves, field and area for 30-day spots.

The observations do, however, indicate a large degree of spasmodic variation of the field. While this may in part be due to difficulties of observation, part of it is undoubtedly genuine. Such variation makes it impossible to extend the analysis to short-lived spots, for which, in any case, the foreshortening correction may be different from that for long-lived spots.

The field shown in Fig. 1 is the (averaged) maximum observable field of a spot. To obtain the total magnetic induction through surface layers of the spot, this maximum field must be multiplied by an effective area of the field; the variation of the latter probably resembles that of the total area of the spot, though not

* These curves are taken from *M.N.*, 85, 553, 1925.

identical with it. Thus the total magnetic induction (unlike the maximum observable field) varies practically in step with the total area; it begins to grow with the area, and fades away with it, while the maxima of the two are not separated by more than about a day. If, then, the magnetic field is produced by the darkening of the spot, a very short time, at most of order one day, is taken to establish the field.

4. We compare this with the time of growth or decay of the sunspot magnetic field, as conditioned by electro-magnetic effects. When a magnetic field decays in conducting material at rest, electric currents are induced, which limit the rate of decay. For simplicity, the field of a sunspot can be taken as axially symmetric, the lines of force lying in planes through the (vertical) axis of symmetry. When such a field decays, no electrostatic fields are set up; thus the induced currents are given by

$$\mathbf{j} = \sigma \mathbf{E} = -\sigma \frac{\partial \mathbf{A}}{\partial t}, \quad (1)$$

where $\mathbf{j}$ is the current-density, $\mathbf{E}$ the electric force, $\mathbf{A}$ the magnetic vector-potential. Again, $\mathbf{j}$ is also connected with the magnetic field by the equation

$$4\pi\mathbf{j} = \text{curl } \mathbf{H} = -\nabla^2 \mathbf{A}. \quad (2)$$

Hence

$$\nabla^2 \mathbf{A} = 4\pi\sigma \frac{\partial \mathbf{A}}{\partial t}. \quad (3)$$

The equation satisfied by $\mathbf{H}$ is less simple in general; but if σ is taken to be constant, it becomes

$$\nabla^2 \mathbf{H} = 4\pi\sigma \frac{\partial \mathbf{H}}{\partial t}. \quad (4)$$

At the surface, $\nabla^2 \mathbf{A}$ is at most of order $\mathbf{A}/a^2$, where a is the radius of the visible spot; it may be smaller, since $\mathbf{A}$ may approximate at the surface to a solution of $\nabla^2 \mathbf{A} = 0$. Hence, by (3), the time of decay of $\mathbf{A}$ is at least of order τ , where

$$\tau = \sigma a^2. \quad (5)$$

This also gives the time of decay of $\mathbf{H}$; for $\mathbf{A}$ is directed round circles with the axis of symmetry of the field as axis, and the total magnetic induction through any such circle equals the line-integral of $\mathbf{A}$ round the circle. In a recent paper* I have found that σ is about 10^{-8} E.M.U. near the Sun's surface. Taking $a = 10^4$ km., the time of decay of the surface field is found to be at least 10^{10} sec., or 300 years—enormously greater than the time in which the sunspot field actually does disappear.

The time of growth of the surface field must have the same order of magnitude. Suppose that, at time $t=0$, currents begin to flow, which give rise to a vector-potential $\mathbf{A}_0$ after a sufficient time has elapsed for the field to attain a constant value. Then the total vector-potential at any time is the sum of $\mathbf{A}_0$ and a part satisfying the induction equations (1-3); that is,

$$\nabla^2 (\mathbf{A} - \mathbf{A}_0) = \sigma \frac{\partial}{\partial t} (\mathbf{A} - \mathbf{A}_0).$$

* Proc. Roy. Soc., A, **183**, 453, 1945.

But at the surface $\nabla^2(\mathbf{A} - \mathbf{A}_0)$ is comparable with $(\mathbf{A} - \mathbf{A}_0)/a^2$; hence $\mathbf{A}$ is comparable with

$$\mathbf{A}_0(1 - e^{-t/\tau}). \quad (6)$$

The time of decay of a magnetic field—*i. e.* the time taken for it to decrease in the ratio $1/e$ —is not the same at all points, nor even at the same point at all times. Suppose that the sunspot field is similar to the field of currents in a solenoid, *i. e.* that, well below the surface, the lines of force form a tight bundle (see Fig. 2A, below). In this bundle, if b is the radius of the bundle's cross-section, $\nabla^2 \mathbf{A}$ is comparable with $\mathbf{A}/b^2$, and so, by (3), the time of decay of the field here is comparable with σb^2 .

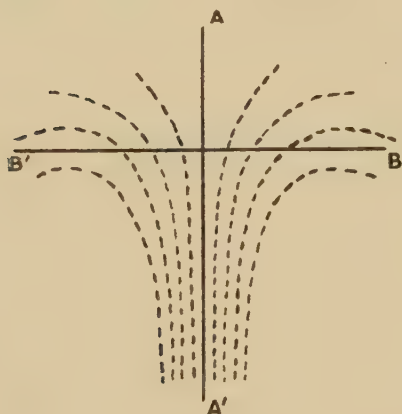


FIG. 2A.

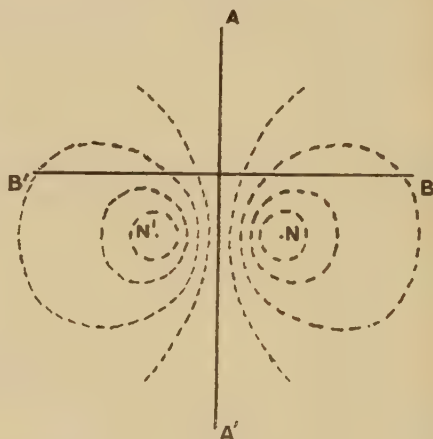


FIG. 2B.

FIG. 2.—Two possible forms for the lines of force in the vertical plane through the axis (AA') of a sunspot. BB' is the Sun's boundary.

This agrees with a result given by Chapman in his interesting and stimulating 1943 Presidential Address to the Society on "Magnetism in the Sun's Atmosphere".* Chapman considered the rate of decay of the magnetic energy in the sunspot field, assumed to be that of a solenoid. If b/a is small, the energy is mainly provided by the field at great depths, and its time of decay is comparable with σb^2 , as Chapman found. But the gulf between the theoretical and observational estimates of the time of decay of the surface field cannot be bridged (as Chapman sought to do) by taking b/a small, and so diminishing the time of decay of the magnetic energy; the theoretical time of decay of the surface field is at least comparable with σa^2 , whatever the time of decay of the magnetic energy.

It may seem surprising that the time of decay of the field can be different at the surface and at great depths. To explain this, we note that (3) and (4) are diffusion equations; they imply that the field is diffusing from point to point. Thus what initially occurs in the solenoidal field of Fig. 2A, if b/a is small, is a swelling of the bundle of lines of force, without any change in the total magnetic induction in the bundle; this hardly affects the surface field. The swelling of the bundle causes both the field on the axis and the magnetic energy per unit length of the bundle to decrease like $1/b^2$.

* *M.N.*, 103, 117, 1943.

5. It follows from the results of the last section that the surface field cannot disappear simply through electro-magnetic decay. This is not surprising as it may appear; electro-magnetic decay is not the only possible cause of its disappearance, nor with an appropriate form for the field, the most probable one. In finding the time of electro-magnetic decay, solar material is supposed to remain still, while the lines of force move through the material; but the material may move, carrying the field with it. The slowness of electro-magnetic decay simply implies that lines of force can move only very slowly through the material; material can move freely along the lines of force, but its motion transverse to the field must be nearly identical with that of the lines of force.

Consider a field like that of Fig. 2A. The mechanical forces due to a magnetic field can be represented by regarding the lines of force as subject to stresses; at a point where the field-strength is H , the stresses consist of a tension $H^2/8\pi$ along the lines of force, and an equal pressure normal to them. At the Sun's surface the lateral pressure is considerably greater than the gas pressure, and causes the lines of force to bend outward. At deep levels, on the other hand, it can be supposed to balance the mechanical pressure-difference between the spot and the surrounding material, due to the coolness of the spot. Lateral radiation tends to warm up the spot material: when this happens, the mechanical pressure-difference between the spot and the surrounding material is reduced, and can no longer balance the magnetic pressure. The bundle of lines of force then expands, and some of the wing lines bend round and disappear below the surface. The partial vacuum due to the expansion sucks up material from great depths along the lines of force; further warming by radiation then occurs, the bundle expands further, and more of the lines of force cease to reach the surface. That is, the warming of the spot is automatically accompanied by a diminution in the spot field. The field at the centre of a spot shows little change during the middle part of the spot's lifetime; this is because the warming, and the consequent expansion of the field, takes place, not through the spot as a whole, but through the region below the penumbra and the edge of the umbra.*

In this, the Sun must be supposed to be stable against convection at levels where the gas pressure is large compared with the magnetic pressure—that is, material pumped up from below must arrive at such levels cooler than the material which it replaces. Otherwise, the pumping results in a further decrease of the pressure-difference between the spot and the surrounding material; the field expands further, more material is pumped up from below, and the process repeats itself. In consequence, in an unstable atmosphere neither the field nor the coolness of a spot could persist for more than at most a few hours when once upward motion had begun. This indicates that the base of a spot must be well below the hydrogen convection zone. The adiabatic cooling of material as it rises must retard the warming of a spot.

The above explanation of the disappearance of the spot field is based on the assumption that the field resembles that shown in Fig. 2A. A second possible form for the field is shown in Fig. 2B; the field here is one due to currents flowing in a ring round the axis of the spot. It is difficult to find a mechanism whereby

* A mechanism like this seems to be assumed by Alfvén in his theory of sunspots (*Ark. Mat. Ast. Fys.*, 29B, No. 2, 1942, and 29A, Nos. 11 and 12, 1943), but he does not set out his ideas in any detail.

the decay of a spot can automatically bring about the decay of this second type of field. To this extent, then, the field shown in Fig. 2A is the more probable in form.

6. As noted in Section 4, the time of 300 years is not only the time of electro-magnetic decay of a sunspot field; it is also the time taken for the field, if due to electric currents, to increase to roughly its steady-state value, against the inertial effect of self-induction. The actual field of a long-lived spot reaches its maximum after (say) 10 days; thus (6) shows that, if it is generated *in situ* by electric currents, the current-producing agency must be 10,000 times as strong as that required to maintain the same field as a steady-state field, and must cease to operate after a few days. This consideration raises serious doubts whether the field actually is generated *in situ*. If the currents are due to the darkness of the spot, they should not cease being generated after a few days. Again, the central magnetic field reaches its maximum at about the same time as the total area, and this maximum is not very sensitive to the size of the area or the details of its growth; there is no reason why this should be so, if the field is generated *in situ*. Finally, it is very difficult to conceive a source of electric currents in the subsurface layers, adequate to produce the observed field as its *steady-state* field; when the currents have to be 10,000 times as great, the difficulty is enormously enhanced. We are therefore led to the view that the field is not generated *in situ*, but that it exists before the spot comes into being, though not in its latter form.

Most probably the growth of the field, like its decay, is due to motion. One possibility is that lines of force of the Sun's general magnetic field are bunched together into the spot as a result of mass motion; this idea forms part of Alfvén's theory. A second possibility is that a field exists below the surface, lines of force running as "girdles" round the Sun's axis, and a pair of spots arises when part of this field is convected to the surface. This alternative would make it possible to incorporate many of the features of the Bjerknes ring-vortex theory* into the magnetic picture. Each of these possibilities is consistent with Fig. 2A, which shows an incomplete field, capable of being linked at great depths to some other field. My own preference is for the second alternative; I hope to set forth my reasons in a later paper.

For the present, however, it is better simply to state the conclusions directly following from the present study. These are that the growth and decay of the sunspot field cannot be due solely to electro-magnetic effects in the absence of convection, but must be due, at least in part, to mass motion; and that the field has the "solenoidal" form shown in Fig. 2A, rather than the "isolated" form of Fig. 2B.

University College of N. Wales,
Bangor :
1946 January 31.

* *Ap. J.*, 64, 93, 1926.

NON-COHERENT SCATTERING AND THE ABSENCE OF POLARIZATION IN FRAUNHOFER LINES

H. Zanstra

(Received 1946 April 8)

Summary

Spitzer has concluded from physical arguments that non-coherent scattering is generally important in stellar atmospheres. It is pointed out that polarization in Fraunhofer lines near the limb would be a measure of coherent scattering, if present, and that the absence of polarization in Redman's observations shows that coherent scattering is largely destroyed.

In a recent paper Spitzer* has concluded that coherent scattering is comparatively rare in stellar atmospheres and that in most cases scattering is largely non-coherent. By this is meant that the re-emitted wave-length is at a different place in the line profile from the absorbed wave-length, so that actually a slight change in wave-length occurs, although the total energy absorbed in the line is equal to that re-emitted. Cases of non-coherent scattering considered by him are: (1) interlocking, (2) Doppler broadening, (3) pressure broadening (hydrogen lines) and (4) impact broadening. In the case of interlocking with broad lower and narrow upper level there is complete redistribution of the re-emitted intensities in the line profile according to statistical weights. The same holds true in case (4) as will presently be seen.

From the observational side Houtgast† has shown that the behaviour of strong Fraunhofer lines across the Sun's disk can only be interpreted under the assumption of non-coherent scattering.

A strong argument that coherent scattering cannot be very important in the formation of Fraunhofer lines is found in the absence of polarization according to Redman's‡ measurements. The author§ had worked out a theory that for coherent scattering from a classical oscillator the corresponding Fraunhofer line would show polarization of its residual intensity for not too large absorption when observed near the limb. For instance the resonance line of neutral calcium $\lambda 4227$, according to Heisenberg||, should behave like a classical oscillator. Then, assuming the Milne-Eddington model, it should show at a residual intensity 0.60 a difference in intensity of the two polarized components of 17.3 per cent. of the continuous spectrum, or 29 per cent. of the intensity at a point close to the limb for which $\sin \theta = 0.985$.¶ The effect of waviness of the solar surface, introduced by Redman in his third eclipse paper**, would cut this polarization down to 0.65 of

* L. Spitzer, *Ap. J.*, **99**, 1, 1944.

† J. Houtgast, *The variations in the profiles of strong Fraunhofer lines along the radius of the solar disc*. Dissertation, Utrecht, 1942.

‡ R. O. Redman, *M.N.*, **101**, 266, 1941; **103**, 229, 1944.

§ H. Zanstra, *M.N.*, **101**, 250, 1941.

|| Cf. Y. Ohman, *M.N.*, **89**, 479, 1929.

¶ The component vibrating parallel to the limb (viz. polarized perpendicular to the limb) should be the strongest, except for the M-E model, when the coefficient of darkening is very small.

** *M.N.*, **103**, 173, 1943.

this or 19 per cent. of the intensity.* In Redman's first observations there was some evidence for a small polarization, but the more accurate observations of the second paper showed that there is no observable polarization and that, if present, it must be on the average less than 0.5 per cent. of the residual intensity. This is only one-fortieth of the theoretical value. Though one should not put too much weight on the actual figures, there is no doubt that by some process or other the coherent scattering is largely destroyed and that one must either have non-coherent scattering or some other process like pure absorption.

Indeed the author† offered an explanation amounting to non-coherent scattering by impact broadening (collision damping), treating the case of a classical oscillator which suffers sudden random changes in phase. If γ_0 is the damping constant for the undisturbed oscillator and γ_1 the collision damping constant, viz. twice the reciprocal of the mean time between collisions, one finds that the oscillator absorbs like an undisturbed oscillator of the same strength f but with damping constant $\gamma = \gamma_1 + \gamma_0$. This was proved if the distance from the line centre is $\gg \gamma_0$, but it is plausible to assume it throughout the line. However, when monochromatic radiation of angular frequency ω is absorbed in the line wing, only the fraction $\gamma_0/(\gamma_1 + \gamma_0)$ of the energy is scattered coherently in the same frequency ω with width zero, whereas the remaining fraction $\gamma_1/(\gamma_1 + \gamma_0)$ of the energy is emitted about the line centre with a frequency distribution which is the same as for the emission of an undistributed oscillator with damping constant $\gamma = \gamma_1 + \gamma_0$.‡ The scattering is therefore non-coherent as far as the second wave is concerned. Moreover, in this second wave there is complete redistribution of the emitted intensities according to a line profile with damping constant $\gamma_1 + \gamma_0$, which is the same as the profile for absorption.

In the case of the *Ca* resonance line $\lambda 4227$ the absence of any appreciable polarization means that the coherent wave is largely suppressed and that one gets a good approximation by considering the second wave only, so that one has scattering with complete redistribution.§ This also follows from the fact demonstrated by Unsöld|| that the total absorption of this Fraunhofer line is abnormally large which he explains by collision damping with a total damping constant $\gamma = 11\gamma_0$, and therefore $\gamma_1 \ll \gamma_0$. According to the above theory the coherent wave is then suppressed.

Redman in addition observed 36 lines in a region about $H\gamma$, 20 Å. wide, and

* The uncorrected $\cos \theta = 0.173$, corrected for waviness, becomes $\cos \bar{\theta} = 0.293$ by interpolation between $\cos \theta = 0.05$ and 0.3 of Redman's $\cos \bar{\theta}$. Then Fig. 4 of the author's paper with $\beta = 4$ yields the correction factor 0.65. It may be remarked that possibly the waviness of the surface is actually greater than that assumed by Redman, which would naturally affect this correction factor. The model used by Redman consists of a Schuster-Schwarzschild scattering atmosphere of optical depth s , surrounding a pure absorption atmosphere with $\eta = 20s$, for τ between 0 and 0.6, and is therefore mainly of the Milne-Eddington type, although Redman looks upon it as an elaboration of the S-S model.

† H. Zanstra, *M.N.*, **101**, 273, 1941.

‡ Between collisions this second wave is $Ae^{-\frac{1}{2}\gamma_0 t} \cos(\omega_0 t + \delta)$ in (16) of the author's paper. The case $\gamma_0 = 0$ has been fully treated (Cf. M. Born, *Optik*, p. 437, Berlin, 1933) and leads to $\gamma = \gamma_1$. It is assumed that the damping of the wave between collisions contributes the term γ_0 . For the present paper, however, this is not of much consequence, since we consider mainly the case $\gamma \ll \gamma_1$.

§ This second wave is re-emitted about the line centre and, if it comes back into the wing of a strong line at all, it can only do so after many scatterings, so that then it is depolarized.

|| *Physik der Sternatmosphären*, p. 276, Berlin, 1938.

also $11\frac{1}{2}$ Å. covering the *Mg* line $\lambda 5184$ and did not detect any polarization. In the former case the difference of the two components was only 0.004 in $\log I$ or 1 per cent. of the residual intensity on the average, in the latter case the difference regardless of sign was 0.006 in $\log I$ or 1.4 per cent., indistinguishable in both cases from observational errors and very much smaller than the polarization expected from a classical coherent scatterer. Although according to Heisenberg not every type of line should show polarization (*Na D*₁ and the *K* line of *Ca*⁺ are important exceptions), its absence near the limb in a great number of Fraunhofer lines shows that as a rule coherent scattering is not very appreciable.

If polarization should be detected in any Fraunhofer line, this would provide the first evidence for coherent scattering on the Sun and enable one to estimate what portion of the absorbed light is scattered coherently. But as long as no polarization is observed, one may reasonably proceed on the assumption that as a rule coherent scattering is practically absent, in line with Spitzer's conclusions which were not based on astrophysical evidence. However, one should perhaps make a reservation for very faint lines which have not yet been accessible to observation of polarization.

In connection with the foregoing some final remarks may be made.

Two processes in which coherent scattering is destroyed may be thought of: non-coherent scattering and pure absorption. In the latter case complete redistribution upon emission takes place but the energy re-emitted is not equal to that absorbed. Of the two processes that of pure absorption seems less likely, except perhaps far into the wings or for weak lines, where the light is coming from a great depth in the atmosphere, when exciting and de-exciting collisions might be important.

Since for the theory the observations should take place near the limb, the absence of polarization, strictly speaking, only shows that coherent scattering is not appreciable for radiation emerging at large angles with the normal to the solar surface. The possibility might therefore be kept in mind that coherent scattering is present in deeper layers and destroyed in the very highest layers, so that not much coherent radiation gets out at large angles. Physically, however, this does not seem very probable, since on account of the lower densities collisions of any kind would become less important in higher layers.

In the case of thermal Doppler broadening only in a pure scattering model, the scattering is coherent for a system moving with the individual atom and should therefore show polarization, if the line depth is not too large so that not too much multiple scattering occurs.

The author wishes to express his appreciation to the Radcliffe Trustees and the Radcliffe Observer, Dr H. Knox-Shaw, for a temporary appointment on the Staff while waiting for a passage to Europe, and to Dr R. O. Redman for valuable advice.

Radcliffe Observatory,

Pretoria:

1946 April.

ON THE GENERAL EQUATIONS OF STELLAR STRUCTURE

R. A. Lyttleton

(Received 1946 March 30)

Summary

The equation of state of gaseous material in convective equilibrium is obtained for the general case when radiation pressure is included without approximation. The equations governing the equilibrium of convective central regions of a star are transformed to a set of equations considerably more convenient for integration purposes. The solutions of these equations, for the distribution of temperature and density, have been derived in the form of power series having the square of the central distance as argument. The first nine terms in each series have been calculated. Successive coefficients are sufficiently regular in trend to enable the question of the range of applicability of the series to be discussed, and it is shown that they will give the values of ρ and T to a high degree of accuracy over a sufficient range to cover any likely size of convective core. The limiting solutions reduce to the Emden functions for $n=3/2$ and 3, and the corresponding series have been calculated. These limiting cases are physically equivalent to zero radiation pressure and zero gas pressure. Intermediate solutions have been made for eleven different values of the central ratio of gas pressure to total pressure. The further problem of fitting on radiative envelopes to these solutions is briefly discussed.

1. *Introduction.*—The object of the investigation described in this paper was to examine the problem of effecting rigorous integrations of the general equilibrium equations of stellar structure when the radiation pressure term is included. Accurate solutions of the full equations are required not only as the first stage in deriving correct theoretical luminosity and radius relations, but also as the only really dependable basis for investigation of such questions as the oscillations and stability of stars and of the prevailing internal conditions. Furthermore, from a mathematical point of view, there is the possibility that important properties of any set of equations may be lost, or spurious properties introduced, when simplifying assumptions are imposed. But the reduction of the general problem to anything approaching manageable form presents considerable difficulty, and only a small amount of progress has been made in this paper.

The starting point adopted is the result established for the case when radiation pressure is negligible, that the central regions of a star must be in convective equilibrium if the energy generation is proportional to a high power of the temperature (or to any function of the temperature increasing with corresponding rapidity). If it is supposed that similar convective conditions hold in the central regions when radiation pressure is not negligible, a first requirement of the theory is to obtain the equation governing convective equilibrium of gaseous material containing appreciable amounts of both thermal and radiant energy. This problem has been solved, and the equation of state appropriate to such gaseous convective matter is shown to be expressible in simple integral form. The resulting relation is then applied to the problem of calculating the density and temperature distribution in the central regions of stars assumed to be convective. The equilibrium equations can be transformed to a fairly convenient form for

integration purposes, but it has not been found possible to solve them other than numerically. The temperature and density, as proportions of their central values, have been derived in the form of power series (in terms of the central distance) for a number of values of the ratio of the gas pressure to total pressure at the centre. The solutions, which can be regarded as generalized Emden functions, are shown to be sufficiently rapidly convergent for them to be used with considerable accuracy for any radius of core likely to be involved in these models. The further question of completing the solutions by fitting on radiative envelopes is not attempted here, but a few remarks on certain aspects of the problem are made at the conclusion of the paper.

2. *The Equation of State of Gaseous Material in Convective Equilibrium and containing Radiant Energy.*—When radiant energy and the pressure due to it are included, the total hydrostatic pressure p in material that would otherwise obey the perfect gas law is

$$p = \frac{\Re}{\mu} \cdot \rho T + \frac{1}{3} a T^4, \quad (1)$$

while E the total energy, thermal plus radiant, per unit volume is

$$E = \frac{\Re}{\mu} \cdot \frac{\rho T}{\Gamma - 1} + a T^4, \quad (2)$$

where $\Re$ = the gas constant,

μ = the mean molecular weight of the material,

Γ = the ordinary ratio of specific heats of the material,

a = Stefan's constant,

and ρ and T denote the density and temperature.

The condition for adiabatic equilibrium to be stable is that

$$\delta(EV) + p\delta V > 0,$$

where the δ here involved refers to a comparison between the energy actually possessed by an element of volume of the material at some given point and the energy it would have if in equilibrium at some near position. When this inequality is satisfied for an element of material, its excess energy causes it to move; that is, convection currents occur. The situation is, however, that even if the energy possessed by the element greatly exceeds the smallest amount necessary for currents to occur, the average temperature and density prevailing at any given point differ only by insignificant amounts from the values they would have if the sign of inequality in the above condition is replaced by that of equality. Thus, when convective equilibrium holds, the equation of state may be derived with sufficient accuracy from the limiting condition

$$\delta(EV) + p\delta V = 0. \quad (3)$$

Since $\delta V/V = -\delta\rho/\rho$, this relation is

$$\frac{\partial E}{\partial T} \delta T = \left(E + p - \rho \frac{\partial E}{\partial \rho} \right) \frac{\delta \rho}{\rho}.$$

Inserting the values of E and p from (1) and (2), the differential form of the equation of state is

$$\rho \left(\frac{\Re}{\mu} \cdot \frac{\rho T}{\Gamma - 1} + 4aT^4 \right) dT - T \left(\frac{\Re}{\mu} \cdot \rho T + \frac{4}{3} a T^4 \right) d\rho = 0. \quad (4)$$

If divided by $\rho^2 T^2$ the left-hand side becomes a perfect differential, thus

$$\frac{4}{3} a d\left(\frac{T^3}{\rho}\right) + \frac{\mathfrak{K}}{\mu} \cdot \frac{1}{\Gamma-1} d(\log T) - \frac{\mathfrak{K}}{\mu} \cdot d(\log \rho) = 0.$$

Writing $(\Gamma-1)^{-1} = n$, the equation of state is therefore

$$\frac{T^3}{\rho} + \frac{3\mathfrak{K}}{4a\mu} \log\left(\frac{T^n \rho_0}{T_0^n \rho}\right) = \frac{T_0^3}{\rho_0},$$

where ρ_0 , T_0 are the values of ρ and T at some definite point in the material. If further, ρ and T are expressed as fractions of ρ_0 and T_0 , which will later be taken as the values at the centre of the star, the equation of state may be written in a standard form, thus

$$\frac{T^3}{\rho} + \frac{1}{k} \log \frac{T^n}{\rho} = 1, \quad (5)$$

where

$$k = \frac{4a\mu T_0^3}{3\mathfrak{K} \rho_0} = \frac{4(1-\beta_0)}{\beta_0}, \quad (6)$$

in terms of the parameter β_0 , the ratio of gas pressure to total pressure at the point where $\rho = \rho_0$, $T = T_0$.

This integral form for the equation of state does not appear to have been obtained hitherto, previous discussions of the effect of radiation pressure being based on the differential form (4). It is of interest to notice that the equation includes, as it should, the usual well-known relations as special cases. Thus, if radiation pressure is negligible, so that β_0 is practically unity, then $k = 0$, and (5) reduces to

$$\log(T^n/\rho) = 0, \quad \text{or} \quad \rho = T^n.$$

While if gas pressure is negligible, so that β_0 is zero, then $1/k = 0$, and (5) becomes

$$\rho = T^3.$$

Equation (5) thus provides the accurate adiabatic relation between ρ and T holding in convective regions of stellar interiors.

3. *The Hydrostatic Equation.*—With the usual assumption of spherical symmetry of the system, the mechanical equilibrium of a layer of material between distances r and $r+dr$ from the centre requires that

$$\frac{d}{dr} \left\{ \frac{\mathfrak{K}}{\mu} \cdot \rho T + \frac{1}{3} a T^4 \right\} = -G\rho r^{-2} \int_0^r 4\pi r^2 \rho \, dr = -G\rho r^{-2} M(r),$$

where G is the constant of gravitation and $M(r)$ the mass within radius r .

This equation can be reduced to a standard form in the usual way by introducing in place of r the dimensionless variable z defined as follows:

$$r = \left\{ \frac{\mathfrak{K}(n+1)T_0}{4\pi\mu G\rho_0} \right\}^{\frac{1}{2}} z,$$

writing $r^{-3}M(r) = 4\pi\rho_0 z^{-3}M(z)$, and then putting $\rho = \rho_0\rho'$, $T = T_0T'$, so that ρ' and T' are the values of the density and temperature expressed as proportions of the central values, ρ_0 and T_0 . Having made these substitutions, the accents attached to ρ' and T' can afterwards be suppressed, but it is to be remembered that ρ and T

will thereafter refer to ratios of the central values of the density and temperature. On this understanding the equation of hydrostatic equilibrium may be written

$$\frac{d}{dz}(\rho T + \frac{1}{4}kT^4) = -(n+1)\rho z^{-2} \int_0^z \rho z^2 dz. \quad (7)$$

The quantity k appearing in the coefficient of T^4 here is exactly the same parameter that has already been defined by equation (6).

It has been found that the values of ρ and T within the (assumed) convective core, as implied by (5) and (7), can be expressed in power series involving only even powers of the distance from the centre. In general, if the variable z is used, the coefficient of the lowest power of z occurring (the term in z^2) in the various series will not be unity. Some simplification is effected in the procedure of solution if a different variable θ given by

$$\theta = \lambda z$$

is adopted, where λ is selected in order that the coefficient of θ^2 in the series for T shall be -1 . Thus, if

$$T = 1 - \theta^2, \quad \text{and} \quad \rho = 1 + b_1 \theta^2,$$

then λ is to be chosen so that these values of T and ρ satisfy (5) and (7) as far as powers of θ up to θ^2 are concerned. From (7) it is found that

$$\lambda^2(-b_1 + 1 + k) = \frac{1}{6}(n+1),$$

and from (5) that

$$b_1 = -\frac{n+3k}{1+k},$$

and hence that the appropriate value for λ is given by

$$\lambda^2 \left[\frac{n+3k}{1+k} + 1 + k \right] = \frac{1}{6} \lambda^2 \alpha = \frac{1}{6} (n+1). \quad (8)$$

Adjusted in this way, the equation of hydrostatic equilibrium becomes

$$\frac{d}{d\theta}(\rho T + \frac{1}{4}kT^4) = -\alpha \rho \theta^{-2} \int_0^\theta \rho \theta^2 d\theta, \quad (9)$$

where α is defined by (8).

4. *Transformation of the Equations for Integration Purposes.*—It would be possible to proceed by solving equations (5) and (9) as they stand by means of series for ρ and T , but it happens that a further important transformation of (9) can be made that considerably simplifies the work. Thus, from (5) it can be shown that

$$\rho^{-1} \cdot \frac{d}{d\theta}(\rho T + \frac{1}{4}kT^4) = (1+n) \frac{dT}{d\theta} + k \frac{d}{d\theta} \left(\frac{T^4}{\rho} \right),$$

and by means of this relation (9) becomes

$$\frac{1}{2}(1+n) \frac{dT}{d\theta} + k \frac{d}{d\theta} \left(\frac{T^4}{\rho} \right) = -\alpha \theta^{-2} \int_0^\theta \rho \theta^2 d\theta.$$

This equation admits of immediate formal integration, with the constant of integration determined at the centre, to

$$(1+n)T + k \left(\frac{T^4}{\rho} \right) = 1 + n + k - \alpha \int_0^\theta \theta^{-2} \int_0^\theta \rho \theta^2 d\theta \cdot d\theta.$$

The special value of this form of the equation is that the last term on the right can immediately be written down when ρ is expressed as a power series in θ , whereas the corresponding term in (9) would involve calculating the product of the ρ series with another series before the second integration could be made.

Finally, by introducing an auxiliary series S defined by $T^3 = S\rho$, the equations to be solved can be put in the integrated form

$$\begin{cases} (I+n)T + kST = I + n + k - \alpha \int_0^\theta \theta^{-2} \int_0^\theta \rho \theta^2 d\theta d\theta, & \text{I} \\ kS + n \log T - \log \rho = k, & \text{II} \\ T^3 = S\rho. & \text{III} \end{cases}$$

These equations are the basis of the numerical integrations, given in this paper, of the original equilibrium conditions. I is a modified integral form of the complete hydrostatic equation. II is the integral form of the condition for convective equilibrium. III is the definition of the simplifying auxiliary function S . The chief advantage of these forms of the equations is that no term in them involves more than a single expansion or modification of any of the literal series that are initially assumed for T , ρ and S . The most awkward operation arising is that of developing the series for $\log T$ from that for T , but, even in this case, once it is done the resulting expression can of course be carried over, simply by a change of letters, to give the required expansion for $\log \rho$. The product ST can be written down at once, as also can $S\rho$, while the term T^3 and the double integral on the right can also be derived without difficulty.

We now put

$$\begin{aligned} T &= a_0 + a_1\theta^2 + a_2\theta^4 + \dots & (a_0 = I, \ a_1 = -I) \\ \rho &= b_0 + b_1\theta^2 + b_2\theta^4 + \dots & (b_0 = I) \\ S &= s_0 + s_1\theta^2 + s_2\theta^4 + \dots & (s_0 = I), \end{aligned}$$

and by means of these express the various terms of I, II and III as power series. The complete coefficients of the powers of θ so obtained can then be equated to zero term by term in the resulting equations. For any particular power of θ , say θ^{2m} , this procedure leads to three equations, linear in the coefficients a_m, b_m, s_m , corresponding to that power, which can be solved to give these coefficients in terms of those of lower order. Since the values of the constant terms are known, the method can be applied to give successively (a_1, b_1, s_1) , (a_2, b_2, s_2) , and so on. Calculations have been made on these lines determining the coefficients up to (a_8, b_8, s_8) .

The equations obtained from I by the various powers of θ are:

$$\begin{aligned} \theta^2: \quad & -\frac{1}{6}\alpha = (I+n)a_1 + k(a_1 + s_1) \\ \theta^4: \quad & -\frac{1}{20}\alpha b_1 = (I+n)a_2 + k(a_2 + a_1s_1 + s_2) \\ \theta^6: \quad & -\frac{1}{42}\alpha b_2 = (I+n)a_3 + k(a_3 + a_2s_1 + a_1s_2 + s_3) \quad \text{etc.*} \end{aligned}$$

* The literal expressions for the equations resulting from equating powers of θ up to θ^{20} for I, II and III, which are not printed in full here, have been deposited with the Society, and are available for reference.

The equations obtained from II are:

$$\theta^2: ks_1 + na_1 = b_1$$

$$\theta^4: ks_2 + n(a_2 - \frac{1}{2}a_1^2) = b_2 - \frac{1}{2}b_1^2$$

$$\theta^6: ks_3 + n(a_3 - a_1a_2 + \frac{1}{3}a_1^3) = b_3 - b_1b_2 + \frac{1}{3}b_1^3 \quad \text{etc.}$$

The equations obtained from III are:

$$\theta^2: 3a_1 = b_1 + s_1$$

$$\theta^4: 3a_2 + 3a_1^2 = b_2 + b_1s_1 + s_2$$

$$\theta^6: 3a_3 + 6a_1a_2 + a_1^3 = b_3 + b_2s_1 + b_1s_2 + s_3 \quad \text{etc.}$$

5. *Numerical Solution of the Equations.*—Apart from the first one or two powers of θ^2 , it is not practicable to determine the coefficients as general parametric functions of n and k . But as the eventual object is to obtain numerical solutions it is convenient at this stage to introduce definite values for these parameters and calculate the coefficients numerically for each particular set of assumed values of n and k .

As far as n is concerned the value 1.5 is adopted throughout on the assumption that the stellar material is at such a stage of ionization that it can be regarded as a monatomic gas. This reduces the set of solutions required for astrophysical purposes to a linear series corresponding to different values of k . The coefficients of the series for T , ρ and S have been computed for thirteen values of k corresponding to the eleven values of β_0 from 1 to 0 at intervals of 0.1 together with two additional values of 0.95 and 0.85 included to cover more closely that part of the range where the variation of solutions with β_0 appeared to be somewhat more rapid than elsewhere. The appropriate values of k and α are given by the relations (6) and (8). The quantities k and α corresponding to the adopted values of β_0 , with $n=1.5$, are given in the following Table I:

β_0	k	α
1.0	0	15
0.95	4/19	7791/437
0.9	4/9	797/39
0.85	12/17	11319/493
0.8	1	51/2
0.7	12/7	4119/133
0.6	8/3	413/11
0.5	4	231/5
0.4	6	411/7
0.3	28/3	2453/31
0.2	16	2031/17
0.1	36	8871/37
0.0	∞	∞

6. *The Solutions for $\beta_0=1$ and $\beta_0=0$.*—These two limiting solutions can seldom if ever be strictly applicable to stellar problems since the former corresponds to complete absence of radiation pressure and the latter to complete absence of gas pressure. Nevertheless they are interesting physically, since they can be used as the simplest approximations to actual conditions, while mathematically they are of interest here as the limiting solutions at the two ends of the k -scale. Their coefficients could also be calculated from the equations obtained here. For

$\beta_0 = 1$, k is zero, and many terms disappear from the coefficient equations. The function S no longer occurs in the first two equations, but as, in this case, it happens to be equal to ρ , the determination of its coefficients involves no extra calculation. For $\beta_0 = 0$, k is infinite, so that only terms involving this factor need be retained, and this again makes for considerable simplification. The function S reduces to unity, so that all the s -coefficients apart from s_0 are zero, and $\rho = T^3$.

It was found that the calculated value of any particular coefficient progresses smoothly with β_0 , and it became evident that an independent calculation of the limiting cases $\beta_0 = 1$ and $\beta_0 = 0$ would afford some check on the general accuracy of expressions used and numerical steps involved in the intermediate cases, which do not lend themselves to any ready means of checking apart from simply repeating the calculations. For the limiting cases, the functions reduce to the ordinary Emden functions of index $3/2$ and of index 3 respectively, as is already well known, so that a check would be provided by the solution in series of the Emden equation. Accordingly it seemed worthwhile to obtain a power series for the particular solution of Emden's equation that starts at unity at the origin and has zero gradient there.

7. *Power Series Solution of Emden's Equation.*—It may easily be shown that for $k=0$ or for $k=\infty$ equations I, II and III reduce to the Emden equation for T , but it is to be noticed that the relations between the independent variable in the resulting equation and the variable z determining the radial distance are different in the two cases. Thus, equation (7) reduces, by means of the relation $\rho = T^n$, to the standard Emden form

$$z^{-2} \frac{d}{dz} \left(z^2 \frac{dT}{dz} \right) + T^n = 0, \quad (10)$$

and when $k=0$, equation (8) shows that $\lambda^2 = 1/6$, so that the relation between z and θ is here

$$\theta^2 = \frac{1}{6} z^2.$$

If $k=\infty$, the equation reduces in just the same way by means of the relation $\rho = T^3$ to the form

$$\theta^{-2} \frac{d}{d\theta} \left(\theta^2 \frac{dT}{d\theta} \right) + 6T^3 = 0.$$

But, while in this case the relation between θ and the variable of the standard Emden equation is as before, the relation between θ and z is different, since now

$$\lambda^2 = \frac{1}{6} (n+1) \left/ \left\{ \frac{n+3k}{1+k} + 1 + k \right\} \right.,$$

and hence for definite θ (that is, the range of the ordinary Emden solution) the value of z is infinite. That is, the radius of a star for which k is infinite would itself be infinite. This is of course a known property of the polytrope $n=3$ if it is regarded as a possible accurate theoretical model of a star.

Thus for both cases the solution in series of powers of θ^2 can be derived from (10), but the interpretation of the variables is different in the two cases. If the equation is written in the form

$$\frac{d^2 y}{dx^2} + \frac{2}{x} \frac{dy}{dx} + y^n = 0$$

the appropriate solution in series can be shown to be

$$\begin{aligned}
 y = & 1 - \frac{1}{3!}x^2 + \frac{1}{5!}nx^4 - \frac{1}{3 \cdot 7!}(8n^2 - 5n)x^6 \\
 & + \frac{1}{9 \cdot 9!}(122n^3 - 183n^2 + 70n)x^8 \\
 & - \frac{1}{45 \cdot 11!}(5032n^4 - 12642n^3 + 10805n^2 - 3150n)x^{10} \\
 & + \frac{1}{135 \cdot 13!}(183616n^5 - 663166n^4 + 915935n^3 - 574850n^2 + 138600n)x^{12} \\
 & - \frac{1}{945 \cdot 15!}(21625216n^6 - 103178392n^5 + 200573786n^4 \\
 & \quad - 199037015n^3 + 101038350n^2 - 21021000n)x^{14} \\
 & + \dots
 \end{aligned}$$

The evaluation of the coefficients in general form as functions of n has not been taken beyond a_7 owing to the great increase in extensiveness of the calculation at each successive stage; the labour of obtaining any one coefficient is much greater than that required for computing all of the preceding coefficients. The values of a_8 that occur in the following tables for $n = 1.5$ and $n = 3$, have been found by using the equations arising from I, II and III when $k = 0$ and $k = \infty$ respectively, so that the coefficients a_8 in the limiting solutions have not been obtained by a strictly independent method.

The values of the coefficients in the series for T , ρ and S , corresponding to the adopted values of β_0 , are set out in Table II.

8. *The Range of Applicability of the Series.*—While it is not possible to derive rigorous values for the range of convergence of these series for T , ρ and S , it is nevertheless a simple matter to obtain trustworthy estimates of the values of θ , and hence of z , within which the series yield a desired accuracy. So long as the value of z is such that the magnitude of successive terms decreases, the series as a whole must converge, since the signs of the coefficients are alternately positive and negative. Also, the error made in stopping the sum at any particular term is less than the value of the next term following it. Thus suppose the value of T is calculated by means of the first eight terms, then the known value of the ninth, viz. $a_8\theta^{16}$, can be used to estimate the accuracy, and it can then be inferred that the accuracy attained by including this term will be at least as good. The values of θ , and the corresponding values of z , at which the ninth term is not greater than 10^{-6} are given in the following Table III:

β_0	θ	z
1.0	0.747	1.83
0.95	0.661	1.77
0.9	0.615	1.76
0.85	0.585	1.77
0.8	0.562	1.80
0.7	0.531	1.87
0.6	0.509	1.97
0.5	0.493	2.12
0.4	0.480	2.32
0.3	0.468	2.63
0.2	0.458	3.17
0.1	0.499	4.40
0.0	0.441	∞

Coefficients in the expansions of T , ρ and S for the selected values of the parameter β_0

β_0	1.0	0.95	0.9	0.85	0.8
a_0	1.0	1.0	1.0	1.0	1.0
a_1	-1.0	-1.0	-1.0	-1.0	-1.0
a_2	0.45	0.43943 0003	0.47199 6689	0.51329 9374	0.55367 6471
a_3	-0.15	-0.15299 9103	-0.20203 0621	-0.25193 4546	-0.29814 9407
T a_4	0.04166 6667	0.05326 5896	0.08634 4569	0.11974 2977	0.15415 4298
a_5	-0.01036 3636	-0.01937 1154	-0.03560 7757	-0.05562 1901	-0.07868 8929
a_6	0.00237 6748	0.00663 4468	0.01448 4901	0.02571 3196	0.03978 4982
a_7	-0.00051 5010	-0.00216 4408	-0.00591 7806	-0.01176 7074	-0.01996 8973
a_8	0.00010 6508	0.00074 1838	0.00238 3181	0.00536 3637	0.00998 5234
b_0	1.0	1.0	1.0	1.0	1.0
b_1	-1.5	-1.76086 9565	-1.96153 8462	-2.12068 9655	-2.25
b_2	1.05	1.57719 1838	2.03479 5052	2.43683 5810	2.79264 7059
ρ b_3	-0.5	-1.05006 5640	-1.59499 1678	-2.13418 9378	-2.65916 3371
b_4	0.19	0.59726 1815	1.07178 6604	1.60767 3236	2.18422 2569
b_5	-0.06179 5454	-0.30642 9592	-0.65213 0998	-1.09935 5811	-1.63167 2728
b_6	0.01802 5350	0.14649 8109	0.37065 0032	0.70307 2109	1.14119 2489
b_7	-0.00482 8372	-0.06652 9920	-0.20033 9791	-0.42802 5952	-0.76040 1195
b_8	0.00121 4085	0.02906 0708	0.10421 7108	0.25095 7637	0.48822 6016
s_0	1.0	1.0	1.0	1.0	1.0
s_1	-1.5	-1.23913 0435	-1.03846 1539	-0.87931 0345	-0.75
S s_2	1.05	0.55915 1100	0.34421 2765	0.23831 7960	0.18088 2353
s_3	-0.5	-0.10657 3122	-0.05483 7339	-0.05327 7134	-0.05587 3085
s_4	0.19	0.00739 8469	0.01945 8025	0.02314 9030	0.02261 1541
s_5	-0.06179 5454	-0.00462 2465	-0.00756 4806	-0.00692 8298	-0.00783 8411
s_6	0.01802 5350	0.00296 0657	0.00172 6240	0.00275 5102	0.00355 1392
s_7	-0.00482 8372	-0.00035 7253	-0.00068 9116	-0.00106 4661	-0.00136 3763
s_8	0.00121 4085	-0.00005 3338	0.00028 1698	0.00038 9184	0.00061 6799

TABLE II—continued

β_0	0.7	0.6	0.5	0.4
a_0	1.0	1.0	1.0	1.0
a_1	-1.0	-1.0	-1.0	-1.0
a_2	0.62375 6237	0.68038 0405	0.72740 2597	0.76800 9832
T a_3	-0.38075 1532	-0.45372 0093	-0.52003 5014	-0.58189 6759
a_4	0.22444 4965	0.29465 0144	0.36439 1450	0.43432 6900
a_5	-0.13102 2000	-0.18949 3347	-0.25308 7919	-0.32180 6260
a_6	0.07573 9710	0.12093 1765	0.17470 4525	0.23727 5820
a_7	-0.04356 5405	-0.07681 9366	-0.12010 9557	-0.17436 8376
a_8	0.02494 9351	0.04862 1604	0.08232 7645	0.12782 6410
b_0	1.0	1.0	1.0	1.0
b_1	-2.44736 8421	-2.59090 9091	-2.7	-2.78571 4286
b_2	3.39412 5830	3.88461 0690	4.29498 7013	4.64638 0158
b_3	-3.64945 6318	-4.55577 3690	-5.38690 5448	-6.15603 6966
ρ b_4	3.40169 0866	4.65181 5043	5.90631 0703	7.15778 7936
b_5	-2.89244 7220	-4.34448 4795	-5.93783 4389	-7.64964 5019
b_6	2.30703 6629	3.81260 4091	5.61853 8522	7.70714 8418
b_7	-1.75543 6746	-3.19574 8285	-5.08384 7042	-7.43403 3380
b_8	1.28833 3999	2.58595 4964	4.44461 0323	6.93431 8645
s_0	1.0	1.0	1.0	1.0
s_1	-0.55263 1579	-0.40909 0909	-0.3	-0.21428 5714
s_2	0.12464 9805	0.09661 3171	0.07722 0779	0.06071 0562
S s_3	-0.05457 0586	-0.04819 4158	-0.04062 2972	-0.03293 7130
s_4	0.02120 1204	0.02045 2808	0.01920 9238	0.01713 1553
s_5	-0.00984 3151	-0.01065 2542	-0.01066 1319	-0.01004 7128
s_6	0.00454 0996	0.00537 7668	0.00590 8766	0.00601 6309
s_7	-0.00218 1235	-0.00289 4900	-0.00343 0608	-0.00373 2028
s_8	0.00107 2866	0.00156 1640	0.00201 6700	0.00235 7119

TABLE II—continued

β_0	0.3	0.2	0.1	0.0
a_0	1.0	1.0	1.0	1.0
a_1	— 1.0	— 1.0	— 1.0	— 1.0
a_2	0.80436 4360	0.83789 4896	0.86955 6077	0.9
a_3	— 0.64102 7252	— 0.69878 5182	— 0.75624 8870	— 0.81428 5714
T a_4	0.50544 2036	0.57882 5833	0.65560 0661	0.73690 4762
a_5	— 0.39629 4429	— 0.47763 6591	— 0.56725 8601	— 0.66689 6104
a_6	0.30957 4587	0.39315 7862	0.49020 0754	0.60354 2957
a_7	— 0.24121 5216	— 0.32305 7985	— 0.42323 2886	— 0.54620 9486
a_8	0.18759 8408	0.26511 5778	0.36517 1559	0.49432 2866
b_0	1.0	1.0	1.0	1.0
b_1	— 2.85483 8710	— 2.91176 4706	— 2.95945 9460	— 3.0
b_2	4.95347 9483	5.22672 1032	5.47368 5931	5.7
b_3	— 6.87595 8189	— 7.55778 4965	— 8.21083 6705	— 8.84285 7143
ρ b_4	8.40837 9075	9.66445 2782	10.93417 9369	12.22642 8571
b_5	— 9.47373 5386	— 11.41421 4598	— 13.48185 6943	— 15.69211 6883
b_6	10.07933 8176	12.75099 9963	15.75050 2408	19.11733 2025
b_7	— 10.27861 0385	— 13.66985 4275	— 17.68126 1497	— 22.40930 3270
b_8	10.14545 2556	14.19776 6957	19.24825 6947	25.49982 5023
s_0	1.0	1.0	1.0	1.0
s_1	— 0.14516 1290	— 0.08823 5294	— 0.04054 0541	zero
s_2	0.04520 1527	0.03004 3240	0.01500 4213	„
s_3	— 0.02521 3815	— 0.01727 9842	— 0.00893 5818	„
S s_4	0.01420 2319	0.01041 8216	0.00572 2076	„
s_5	— 0.00879 1505	— 0.00680 2020	— 0.00393 6751	„
s_6	0.00562 0402	0.00461 0042	0.00281 7465	„
s_7	— 0.00370 4944	— 0.00321 6291	— 0.00207 4514	„
s_8	0.00249 1375	0.00229 0260	0.00155 9233	„

Values of T and ρ satisfying I, II and III are required for the range of values of z determined by the radius of the convective core, and this distance itself depends on the surrounding envelope, in that the change from convective to radiative equilibrium must occur at such a value of z that the resulting subsequent values of T and ρ shall run out to zero together, thereby leading to proper surface conditions for the configuration. For $\beta_0=1.0$ the appropriate value is known to be about $z=1.19$, and the series would therefore give an accuracy much closer than 10^{-6} at all parts of the core. The radius of the core, as measured by z , may be expected to increase with decreasing β_0 , but Table III shows that the limiting value of z for a given accuracy also increases, and it may reasonably be expected that the series will continue to be capable of giving comparable accuracy for all values of β_0 . If it should be found that accuracy of only 10^{-4} were sufficient, the range of z would be increased by just about one-third.

It will be seen that the series for S is more rapidly convergent than that for T for all the intermediate values of β_0 , so that S can be calculated to higher accuracy than T . The series for ρ , however, is somewhat less convergent at corresponding values of θ , so that in any instance in which the ρ series might not give adequate accuracy, it would be possible to proceed otherwise and obtain ρ directly from II or III, making use of the more accurately determinable values of T and S .

9. *The Radiative Envelopes.*—There remains to be carried out the heavy task of completing the configurations by determining the appropriate radiative envelopes. As already mentioned, the boundary conditions to be satisfied by the envelope are that it shall have temperature and pressure continuity with the convective core,

and itself be just about to become radiatively unstable at its inner boundary; that is, equation (3) must be satisfied on the radiative side of the interface. Also, the value of z at which the type of equilibrium changes must be such that the values of ρ and T in the envelope tend to zero together at the surface. Thus, mathematically, the most direct plan for calculating the values of ρ and T in the envelope would be to adopt in turn a number of values for the radius of the core and attempt to fit on envelopes satisfying these conditions. The direction in which any solution so attempted fails to fulfil the external boundary condition would then serve as a guide to a closer estimate of the core radius. Proceeding in this way the correct value would, in principle, gradually be approached. It is found, however, that this procedure is barely practicable even when radiation pressure is neglected, for the conditions in the outer part of the envelope are so highly sensitive to the adopted radius of the core that solutions of great accuracy are needed in order to permit a sufficient range in z to be covered. Thus each of the trial solutions is very laborious, and it is a further disadvantage that the unsuccessful integrations cannot be made use of in any of the subsequent attempts.

The process of inward integration, on the other hand, does not seem to involve this difficulty concerning accuracy, the integrations, if anything, tending to converge towards the true solution in spite of any small unavoidable errors. But it differs from outward integration in that trial numerical values have to be adopted for the radius and interior mass, or assumptions equivalent to this made. Thus the linear series of solutions necessitating high numerical accuracy for their derivation by outward integration is replaced for inward integration by a double series of solutions obtainable with some relaxation in accuracy. The problem of overcoming the difficulty presented by this appearance of a doubly infinite set of starting values has recently been discussed by Hoyle*, who has shown that when radiation pressure is negligible it is possible, with only slight approximation, to reduce the inward solutions to a linear series. The process suggested, however, leads to an equation of not inconsiderable complexity for the condition of fit at the limit of the convective core, and this equation itself would have to be solved by recourse to some trial and error method. Nevertheless it seems likely that the method of inward integration would involve less labour than the direct method of outward integration.

It is important to notice that the parameter k , defined by (6), corresponding to any particular star of given mass and composition, cannot be fixed beforehand and the appropriate solution for the convective core selected. It is only determined subsequently by the condition that the total mass shall have the proper value. This feature differentiates the solution of the general problem from the solutions obtained from simpler models forming homologous series.

A further complication that would need to be considered in deriving the envelope solutions, especially for intermediate values of β_0 (that is, for stars of intermediate mass) is the additional contribution to the opacity by electron scattering. This feature would lead to the occurrence of a second parameter, similar to k , in the equation of radiative equilibrium, which would also have to be determined subsequently when the mass and composition were specified.

When the work described in this paper was begun, I was unaware that the

* *M.N.*, 105, 23, 1945.

problem was already under examination by Henrich.* The method adopted in his paper, however, differs considerably from the present method. The equations arrived at are also different both in form and appearance, but it has been verified that the two systems of equations are equivalent. In Henrich's work it is found possible to express the pressure, density and temperature at all parts of the convective region in terms of the ratio of radiation pressure to the gas pressure, and this latter ratio, or rather a certain function of it, is used as the dependent variable in effecting the integrations. The integral form of the equation of state, which forms the basis of the calculations of the present paper, is not explicitly obtained by Henrich. The particular cases selected by him for numerical integration correspond to values of k equal to 0.04, 0.4, 1, 2, 4 and 8, so that only two of the solutions given in this paper are for cases for which solutions are already available. The two papers represent substantially different modes of attack on the problem, so that there has been no duplication of the details of the theoretical part of the study.

Cambridge :

1946 March.

* *Ap. J.*, **93**, 483, 1941.

ERRATUM

M.N., 106, No. 2, M. Minnaert's paper :

P. 100, last line but two, *for* latitude *read* altitude.

